(NASA-CR-155081) RESEARCH IN THE SPACE
SCIENCES. VOLUME 2, NO. 1: JANUARY - MARCH
1975 (California Univ.) 125 p

N77-86304
THRU
N77-86313
Unclas
00/99 51184

RESEARCH IN THE SPACE SCIENCES

A Bibliography of Research Performed
Through the Space Sciences Laboratory,
University of California, Berkeley
with partial support from
NASA Grant NGL 05-003-497

Kinsey A. Anderson, Director
and
Principal Investigator

Volume II, No. 1

January - March, 1975

Space Sciences Laboratory, Series 16, Issue 26

TABLE OF CONTENTS

C.S. Bowyer and Group (Cont.)

Soft X-ray Observations of Centaurus X-3 from Copernicus, B. Margon, K.O. Mason, F.J. Hawkins and P.W. Sanford, Ap. J. (Letters), 196, L51, 1975

A.L. Burlingame and Group

Ketones Derived from the Oxidative Degradation of Green River Formation Oil Shale Kerogen, B.R.T. Simoneit and A.L. Burlingame, Advances in Organic Geochemistry, 1973, 6th International Congress of Organic Geochemistry

The Maturation of Lunar Dust Grains, J.L. Bertaut, J.P. Bibring, J. Borg, A.L. Burlingame, Y. Langevin, M. Maurette, F.C. Walls, Lunar Science VI, The Lunar Science Institute, Houston

Simulation of Lunar Carbon Chemistry: I. Solar Wind Contribution, J.P. Bibring, A.L. Burlingame, J. Chaumont, Y. Langevin, M. Maurette and P.C. Wszolek, Proceedings of the Fifth Lunar Conference, 2, 1747-1762, 1974

Study of the Organic Matter in the DSDP (Joides) Cores, Legs 10-15, B.R.T. Simoneit and A.L. Burlingame, Advances in Organic Geochemistry, 1973, 6th International Congress of Organic Geochemistry

Unimolecular Gas-Phase Hydrogen Randomization within 2-Methylpropane Radical Cations, P.J. Derrick, A.M. Falick and A.L. Burlingame, J. Chem. Soc., 98-101, 1975

M. Calvin and Group

Kerogen Structures in Recently-Deposited Algal Mats at Laguna Mormona, Baja California: A Model System for the Determination of Kerogen Structures in Ancient Sediments, R.P. Philp and M. Calvin, 2nd Environmental Bio-Chemistry Conference, Burlington, Canada, 1975

C.W. Churchman and Group

The Costs of Efficiency, I.R. Hoos, JHE, XLVI, 2, 1975

Author's Reply to Technology Transfer: Another Opinion, I.R. Hoos, J. Dynamic Systems, Measurement, and Control, 1975

B. Fraenkel-Conrat

The Specificity of Different Classes of Ethylating Agents toward Various Sites in RNA, B. Singer and H. Fraenkel-Conrat, Biochemistry, 14, 772, 1975

B. Fraenkel-Conrat (Cont.)

The Chemical Effects of Nucleic Acid Alkylation and their Relation to Mutagenesis and Carcinogenesis, B. Fraenkel-Conrat, Progress in Nucleic Acid Research and Molecular Biology, 15, 219-332, 1975 (reprint not available)

H.H. Heckman and Group

Isotope Production Cross Sections from the Fragmentation of ^{16}O and ^{12}C at Relativistic Energies, P.J. Lindstrom, D.E. Greiner, H.H. Heckman, B. Cork and F.S. Bieser, Phys. Rev. Lett., 1975 (in press)

Momentum Distributions of Isotopes Produced by Fragmentation of Relativistic ^{12}C and ^{16}O Projectiles, D.E. Greiner, P.J. Lindstrom, H.H. Heckman, B. Cork and F.S. Bieser, Phys. Rev. Lett., 1975 (in press)

F.S. Mozer and Group

Properties of the Wake of Small Langmuir Probes on Sounding Rockets, E.A. Bering, J. Atmos. and Terr. Phys., 37, 119-129, 1975

The Temperature Gradient Drift Instability at the Equartorward Edge of the Ionospheric Plasma Trough, M.K. Hudson and M.C. Kelley, 1975 (submitted)

Finite Heat Conduction Effects on Ion Cyclotron and Drift Cyclotron Modes, M.K. Hudson, 1975 (in press)

P.B. Price and Group

Composition and Energy Spectra of Heavy Nuclei of Unknown Origin Detected on Skylab, J.H. Chan and P.B. Price, Phys. Rev. Lett., 1975 (submitted)

Micrometeorites and Solar Flare Particles In and Out of the Ecliptic, I.D. Hutcheon, J. of Geophys. Res., 1975 (in press)

Plutonium-224 Fission Tracks: An Alternative Explanation for Excess Tracks in Lunar Whitlockites, I.D. Hutcheon and P.B. Price, Science, 187, 864, 1975

Slowly Braked, Rotating Neutron Stars, H. Sato, Astrophys. J., 195, 743-749, 1975

Towards Diffraction-Limited Seeing of Dim Astronomical Objects from the Earth's Surface using Lidar, B.G. Cartwright and A.J. Gadgil, Astrophys. J., 1975 (submitted)

C.H. Townes and Group

The Methanol Source in Orion at 1.2 Centimeters, M.F. Chui, A.C. Cheung, D. Matsakis, C.H. Townes and A.G. Cardiasmenos, Astrophys. J., 187, L19-L21, 1974

Molecule Searches in Comet Kahoutek (1973f) at Microwave Frequencies, S.A. Mango, K.J. Johnston, J.F. Chui, A.C. Cheung and D. Matsakis, Icarus, 23, 590-594, 1974

Observations and Analysis of the Jovian Spectrum in the 10 Micron $_2$ Band of the NH_3, J.H. Lacy, A.I. Larrabee, E.R. Wollman, T.R. Geballe, C.H. Townes, J.D. Bregman and D.M. Rank, Ap. J. Lett., 1975 (submitted)

Spectroscopy of the Orion Nebula from 80 to 135 Microns, D. Brandshaft, R.A. McLaren and M.W. Werner, Ap. J. Lett., 1975 (submitted)

J.D. Young

The Encephalitogenic Dose Response in Guinea Pigs of the Tryptophan Region of Myelin Basic Protein, J.D. Young, D. Tsuchiya, M. Geier, S. Geier, F.D. Westall, M. Thompson, R. Cyr, E. Ward and F. Yurochko, Immunological Communications, 3(3), 219-226, 1974

Experimental Allergic Encephalitis: Study of Cellular Immunity to the Enphalitogenic Determinant, L.E. Spitler, C.M. von Muller, J.D. Young, Cellular Immunology, 15, 143-151, 1975

The Encephalitogenic Dose Response in Guinea Pigs of the Tryptophan Region of Myelin Basic Protein, J.D. Young and F.C. Westall, South African Journal of Science, 70, 252, 1974

A Synthetic Peptide as a Substrate for Cyclic-Amp-Dependent Protein Kinase, P. Daile, P.R. Carnegie and J.D. Young, Proc. Aust. Biochem. Socl, 8, 1975

THE FLASH PHASE OF SOLAR FLARES: SATELLITE OBSERVATIONS OF ELECTRONS

R. P. LIN

Space Sciences Laboratory, University of California, Berkeley, Calif. 94720, U.S.A.

Abstract. Satellite observations of solar electrons bearing on flare particle acceleration and the generation of radio and X-ray emission are reviewed. The observations support a two stage acceleration process for electrons, one stage commonly occurring at the flare flash phase and accelerating electrons up to ~ 100 keV, and a second stage occurring only in large proton flares and accelerating electrons up to relativistic energies. The location of the acceleration region appears to be no lower than the lower corona.

The accelerated non-relativistic electrons generate type III radio burst emission as they escape from the Sun. Direct spacecraft observations of the type III emission generated near 1 AU and the energetic electrons, provide quantitative information on the characteristics of the electrons exciting type III emission, the production of plasma waves, and the conversion from plasma waves to electromagnetic radiation.

1. Introduction

In the past decade spacecraft penetrating beyond the Earth's magnetosphere have been able to directly observe the low energy particle emissions of the Sun. Low energy solar particles are very frequently emitted from the Sun, especially non-relativistic electrons (see Table I) which commonly originate in small importance 1 flares or subflares. Electrons up to $\sim 10^8$ eV in energy have been observed (Datlowe, 1971) from larger flares. Energetic electrons appear to contain the bulk of the flare energy

TABLE I

Solar particle events in an active year

Number of solar flares[a]	$\sim$16,000
Number of non-relativistic electron events[b]	$\sim$ 400
Number of energetic proton events[b]	$\sim$ 70

[a] Normalized for the whole Sun = number observed × 2.
[b] Normalized by cone of emission of $\sim 70°$ for non-relativistic electron and $\sim 100°$ for proton events.

in those flares where they are accelerated (Lin and Hudson, 1971; Syrotvatskii and Shmeleva, 1972) and they are responsible for most of the observed flare energetic X-ray and radio emission. The particle observations, combined with improved observations of solar electromagnetic radiation (much of which has also been provided by spacecraft) define a detailed physical picture of energetic electrons in small solar flares. Here we review the results obtained from direct spacecraft sampling of electrons from flares pertaining to the acceleration of electrons at the sun and to the generation of type III radio emission at 1 AU.

Gordon Newkirk, Jr. (ed.), Coronal Disturbances, 201–223.

2. Interplanetary Propagation

The electron observations at 1 AU are for the most part limited to the particle intensity above a given energy threshold. The intensity vs time profiles for flare associated events are generally consistent with an impulsive injection into the interplanetary medium. If the electrons are scattered a great deal in their propagation from the Sun to 1 AU (mean free path $= \lambda \ll 1$ AU) then the intensity-time profile will look diffusive, as in Figure 1. If, on the other hand, very little scattering occurs ($\lambda \gtrsim 1$ AU) then the

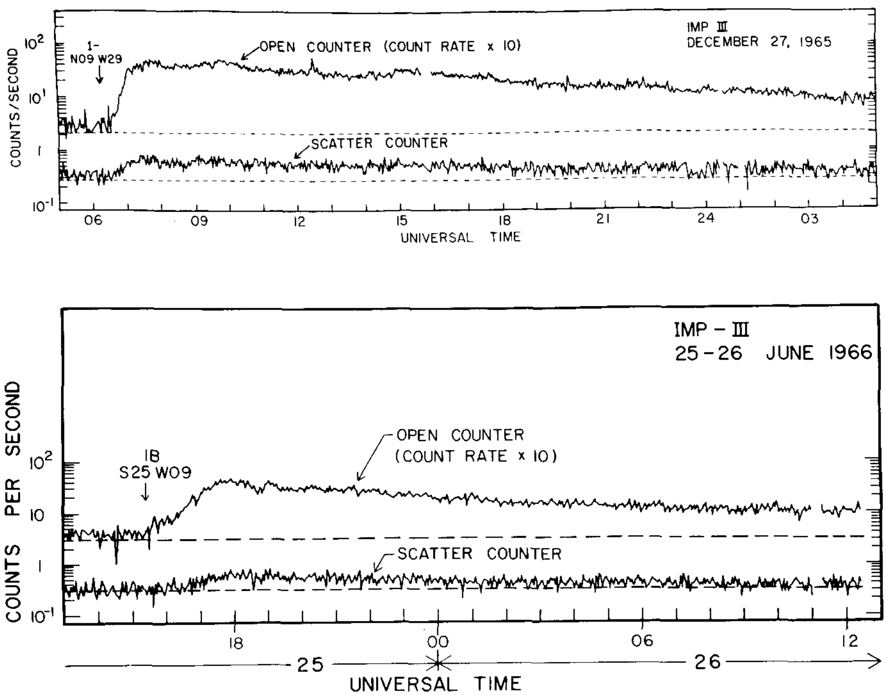


Fig. 1. Two diffusive electron events. The scatter counter is only sensitive to >45 keV electrons while the open counter counts both >40 keV electrons and >0.5 MeV protons. The 27 December 1965 event shows a rapid rise during onset which is observed for many electron events.

intensity-time profile will consist of a rapid increase and decrease as in Figure 2. Bot kinds of events, diffusive and scatter-free (Lin, 1970), and events in between as well are observed, with scatter-free events ($\lambda > 1$ AU) numbering $\sim 20\%$ of the total.

It might seem, then, that it would be very difficult to derive the characteristics of th injected electrons since the observations at 1 AU are greatly affected by the amoun of scattering and how it varies with energy, etc. Actually, however, the *maximum flu* is remarkably insensitive to the details of the scattering (Lin, 1971), so long as there i enough scattering to give diffusive profiles. We can obtain the relationship betwee the maximum flux and the number of particles emitted for an event which can b

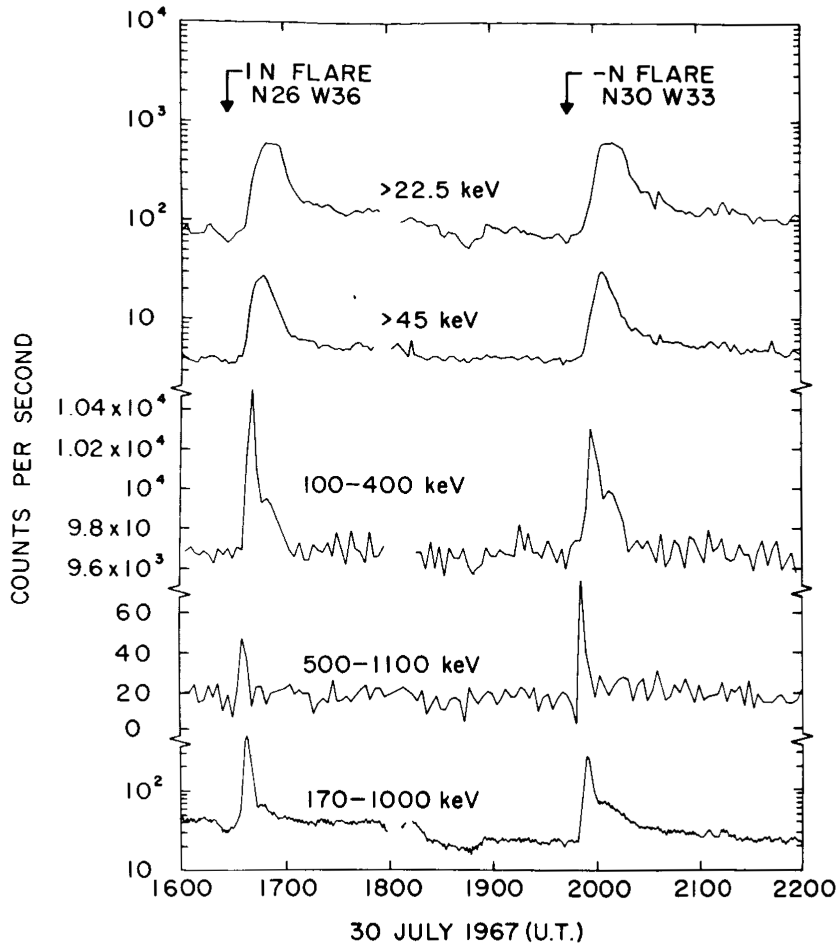


Fig. 2. Two scatter-free impulsive solar electron events from McMath plage 8905 (Wang *et al.*, 1971). The sharp initial peak in the high energy channels place an upper limit on the duration of the electron injection into the interplanetary medium of $\lesssim 3$ min. Note the long decay following the initial peak in the low energy channels.

described by a diffusion equation (Parker, 1963)

$$\frac{\partial \varrho}{\partial t} = \frac{1}{r^{\alpha}} \frac{\partial}{\partial r} \left(M r^{\alpha+\beta} \frac{\partial \varrho}{\partial r} \right), \tag{1}$$

where the diffusion coefficient $D = Mr^{\beta}$, M and β are parameters. Here ϱ is the particle density at position r, time t, of energy E and α is the parameter specifying the dimension of the space to be used. From Krimigis (1965), we obtain for an assumed isotropic flux $J = \varrho v / 4\pi$,

$$J(r, t) = \frac{Nv \exp\left[-\frac{1}{Mt} \frac{r^{2-\beta}}{(2-\beta)^2} \right]}{4\pi (2-\beta)^{(2\alpha+\beta)/(2-\beta)} \, \Gamma[\alpha+1)/(2-\beta)]} \left(\frac{1}{Mt} \right)^{(\alpha+1)/(2-\beta)}, \tag{2}$$

where N = number of particles emitted per unit solid angle, and v = particle velocity. At time of maximum we obtain

$$J_{\max}(r, t_{\max}) = \frac{Nv \exp\left[-\dfrac{\alpha-1}{2-\beta}\right]}{4\pi(2-\beta)^{(2\alpha+\beta)/(2-\beta)}\,\Gamma[(\alpha+1)/(2-\beta)]}\left(\frac{(2-\beta)(\alpha+1)}{r^{2-\beta}}\right)^{(\alpha+1)/(2-\beta)}. \quad (3)$$

Note that the relationship between N and $J_{\max}$ is independent of M. Thus, as long as the diffusion coefficients for the different energy particles have the same spatial dependence (same β) regardless of the value of M, the constant of proportionality between $J_{\max}$ and Nv remains the same. In this case the shape of the flux spectrum derived from the maximum flux at each energy observed at 1 AU is exactly that of the emitted particles.

Some computations have been made of the validity of this method as λ increases (Lin *et al.*, 1973b). These indicate that up to $\lambda \lesssim 0.3$ AU this constant relationship holds.

3. Location of the Acceleration Region

The electrons escaping to the interplanetary medium will lose energy during their passage through the solar atmosphere overlying the acceleration region. A straightforward calculation, assuming rectilinear upward path through fully ionized hydrogen to 1 AU without regard to deflections or other energy loss mechanisms (such as wave-particle interactions or radio emission), (Lin, 1973) shows that an initial power law spectrum at height h at the Sun

$$\frac{\mathrm{d}n}{\mathrm{d}E_1} = AE_1^{-\delta} \quad \text{with } A, \delta \text{ constants} \quad (4)$$

becomes a peaked spectrum

$$\frac{\mathrm{d}u}{\mathrm{d}E_2} = \frac{AE_2}{(E_2^2 + 2k)^{(\delta+1)/2}} \quad (5)$$

with peak at

$$E_{2m} = \left(\frac{2k}{\delta}\right)^{1/2} \quad (6)$$

and $k = -2.6 \times 10^{-18} \int_h^{1\,\mathrm{AU}} n_i(x)\,\mathrm{d}x$, where n_i is in cm^{-3}, h in cm.

In actuality the helical paths of the electrons along the field line and deflections will increase the path length traversed by the electrons in escaping to 1 AU, so that the E_{2m} given by Equation 6 will be a lower limit. A few low energy spectra of electrons observed at 1 AU have become available from recent spacecraft observations (Figure 3). All of these spectra extend smoothly in a power law to below ~6 keV, and on occasion to lower energies. Using Equation (6) we find that a peak $\lesssim 6$ keV implies that the total path length is given by

$$\int_{h}^{1\,\mathrm{AU}} n_i(x)\,\mathrm{d}x \lesssim 3.5 \times 10^{19}\ \mathrm{cm}^{-2}.$$

This corresponds to $\lesssim 60\ \mu\mathrm{g\ cm}^{-2}$ of hydrogen, equivalent to an ambient density at the acceleration region of certainly less than $\sim 10^{10}\ \mathrm{cm}^{-3}$. This density corresponds to a height of $\gtrsim 2 \times 10^4$ km above the photosphere for a 10 X Baumbach-Allen active region density model.

We wish to re-emphasize the fact that this estimater is a *lower* limit to the actual

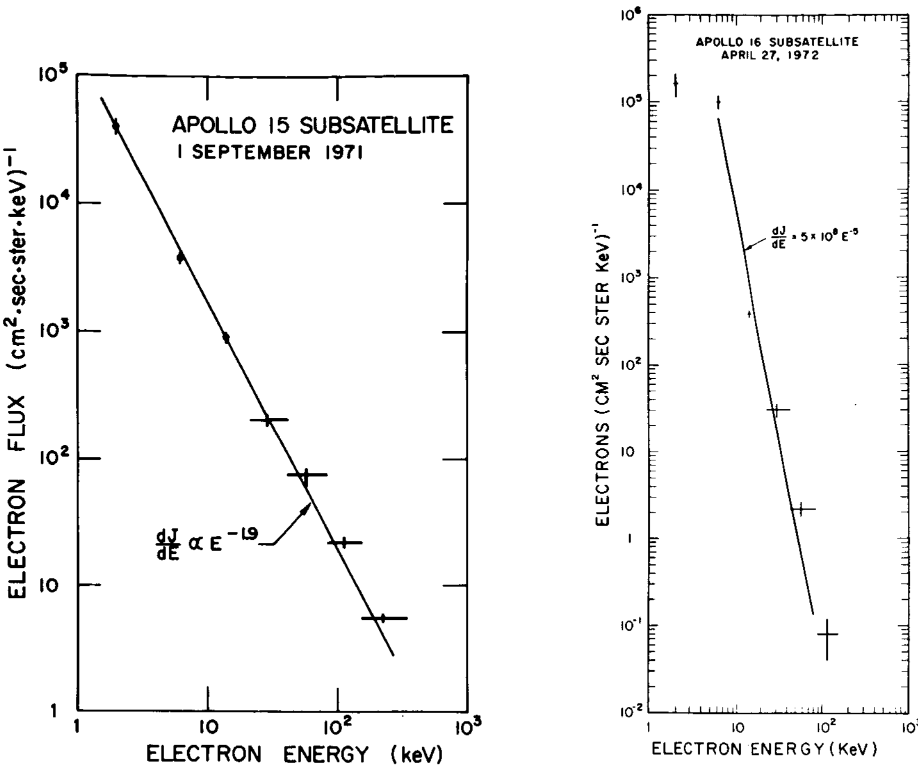


Fig. 3. Two electron energy spectra extending to low energies. The 1 September event is accompanied by energetic protons while the 27 April event (from Lin *et al.*, 1973) is not. Both spectra extend smoothly in a power law to below ~ 6 keV.

height of acceleration since the effects which were not taken into account would tend to increase the minimum energy of the peak. Clearly electron acceleration must have occurred in the lower corona. Although only a few events have been observed to energies below ~ 20 keV, in no events has a turnover been observed at higher energies. Thus electron acceleration at the flash phase appears to be a coronal phenomenon, at least for events observed to emit electrons into the interplanetary medium.

This location and ambient density is consistent with the observed starting frequencies (~200–1000 MHz) of type III bursts, and is also consistent with the occasional observation of an electron event at 1 AU without detectable X-ray emission (Kane and Lin, 1972). Presumably in those events the magnetic field structure in the vicinity of the acceleration region is such as to prevent the electrons from entering dense regions where a detectable X-ray flux would be produced.

4. Accelerated vs Escaping Electron Energy Spectra

Non-thermal X-rays give direct information on the electrons at the Sun. The X-ray spectrum can be directly related to the instantaneous X-ray producing electron spectrum (Brown, 1971; Kane and Anderson, 1970). This relationship can be written for a power law X-ray spectrum as

$$\frac{\mathrm{d}n}{dE} = 3.85 \times 10^{41}\ \gamma(\gamma-1)^2\ B(\gamma-\tfrac{1}{2}, \tfrac{3}{2})\ \frac{AE^{-\gamma+1/2}}{n_i V}\ \mathrm{cm}^{-3}\ \mathrm{keV}^{-1}, \tag{7}$$

where

$$\frac{\mathrm{d}J(h\nu)}{\mathrm{d}(h\nu)} = A(h\nu)^{-\gamma}.$$

$B(x, y)$ is the beta function, n_i the ambient ion density, V the volume of the X-ray region, and A, γ constants.

The relationship of the instantaneous X-ray producing electron spectrom to the accelerated electron spectrum depends on the evolution of the electrons subsequent to acceleration. Suppose the electrons are accelerated in one region and produce the bulk of the observed X-rays in another region (these two regions may be one and the same but for the sake of generality we will allow them to be different). The evolution of the electron distribution, $N(E, t) = V(\mathrm{d}n/\mathrm{d}E)$, where $V =$ volume in the X-ray emitting region, can be described by the equation

$$\frac{\partial N(E, t)}{\partial t} = F(E, t) - \frac{N(E, t)}{\tau_e(E)} - \frac{\partial}{\partial E}\left[N(E, t)\frac{\mathrm{d}E}{\mathrm{d}t}\right], \tag{8}$$

where $F(E, t)$ is the input source of electrons $\mathrm{keV}^{-1}\ \mathrm{s}^{-1}$, $N(E, t)/\tau_e(E)$ is the number of electrons escaping the region $\mathrm{s}^{-1}\ \mathrm{keV}^{-1}$, and the third term describes energy loss processes for the electrons. Note that X-ray observations define $N(E, t)$ subject to a choice of ambient density n_i (Equation (7)). Thus this equation can be solved for $F(E, t)$, given n_i and given the form of $\tau_e(E)$ if only collisional energy losses are assumed to be important.

To a good approximation we can consider $N(E, t)$ as constant over some time interval, Δt, and zero outside that interval. This removes the time dependence of the equation. Additionally we shall consider only the power law case, $N(E, t) = BE^{-\delta}$, so that inserting for $\mathrm{d}E/\mathrm{d}t$ the energy loss in ionized hydrogen (Trubnikov, 1965)

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -4.9\times10^{-9}\, n_i E^{-1/2} (\mathrm{keV\ s^{-1}}), \tag{9}$$

where n_i = ambient density in cm^{-3} and E is in keV, Equation (8) becomes

$$F(E) = N(E)\left[\frac{1}{\tau_e(E)} + \frac{4.9\times10^{-9}\, n_i(\delta+\frac{1}{2})}{E^{3/2}}\right]. \tag{10}$$

We have computed the anticipated energy dependence of $F(E, t)$ compared to $N(E, t)$ and $\mathrm{d}J(h\nu)/\mathrm{d}(h\nu)$, and the energy dependence of the escaping electrons for two extremes:

(1) where the escape term is much larger than the collisional energy loss term. This situation is the *thin-target* approximation for X-ray emission.

(2) where the collisional energy loss term is much larger than the escape term. This situation is the *thick-target* approximation for X-ray emission.

We have used two obvious choices for the energy dependence of τ_e, although other forms might be appropriate. These two are: (1) τ_e = constant, and (2) $\tau_e \propto 1/E^{1/2}$, i.e., proportional to the scale size of the X-ray region divided by the particle velocity. The results are summarized in Table II.

TABLE II

Spectral dependence of electrons and X-rays

	Thick target	Thin target
Spectrum of X-rays	$\frac{\mathrm{d}J(h\nu)}{\mathrm{d}(h\nu)} = A(h\nu)^{-\gamma}$	$\frac{\mathrm{d}J(h\nu)}{\mathrm{d}(h\nu)} = A(h\nu)^{-\gamma}$
Spectrum of electrons in X-ray emitting region $N(E) \propto \frac{\mathrm{d}ne}{\mathrm{d}E} \propto E^{-\delta}$	$\delta = \gamma - \frac{1}{2}$	$\delta = \gamma - \frac{1}{2}$
Spectrum of accelerated electrons $F(E) \propto E^{-\delta_a}$	$\delta_a = \gamma + 1$	$\delta_a = \gamma - \frac{1}{2}$ for τ_e = constant $\delta_a = \gamma - 1$ for $\tau_e \propto E^{-1/2}$
Spectrum of electrons escaping from the X-ray region, $S(E) \propto E^{-\delta_e}$	$\delta_e = \gamma - \frac{1}{2}$ for τ_e = constant $\delta_e = \gamma - 1$ for $\tau_e \propto E^{-1/2}$	$\delta_e = \delta_a =$ $\gamma - \frac{1}{2}$ for τ_e = constant $\gamma - 1$ for $\tau_e \propto E^{-1/2}$

The spectrum of electrons escaping from the X-ray region is not necessarily the spectrum of the electrons escaping to the interplanetary medium. The electrons need not escape to the interplanetary medium to be lost from the X-ray region; they may also escape to the low density, $n_i \lesssim 10^9\ \mathrm{cm}^{-3}$, upper corona, where the flux of X-rays they produce will be below the threshold of current X-ray detectors. Also the acceleration region may be much higher in the solar atmosphere than the X-ray region, and the electrons observed in space may have come directly from the accelerated population (i.e., $\delta_e = \delta_a$).

Datlowe and Lin (1973) noted that it is possible to distinguish between thick and thin target cases under the assumption that the spectrum of electrons observed in the interplanetary medium is representative of the accelerated electron spectrum (i.e., $\delta_e = \delta_a$, see Table II). For a flare event where high energy resolution measurements were available for the electrons and X-rays above 20 keV (see Figure 4), the result was $\delta_a = \gamma - \frac{1}{2}$, favoring thin target. Other X-ray electron events studied where only measurements with poor energy resolution were available are also generally consistent

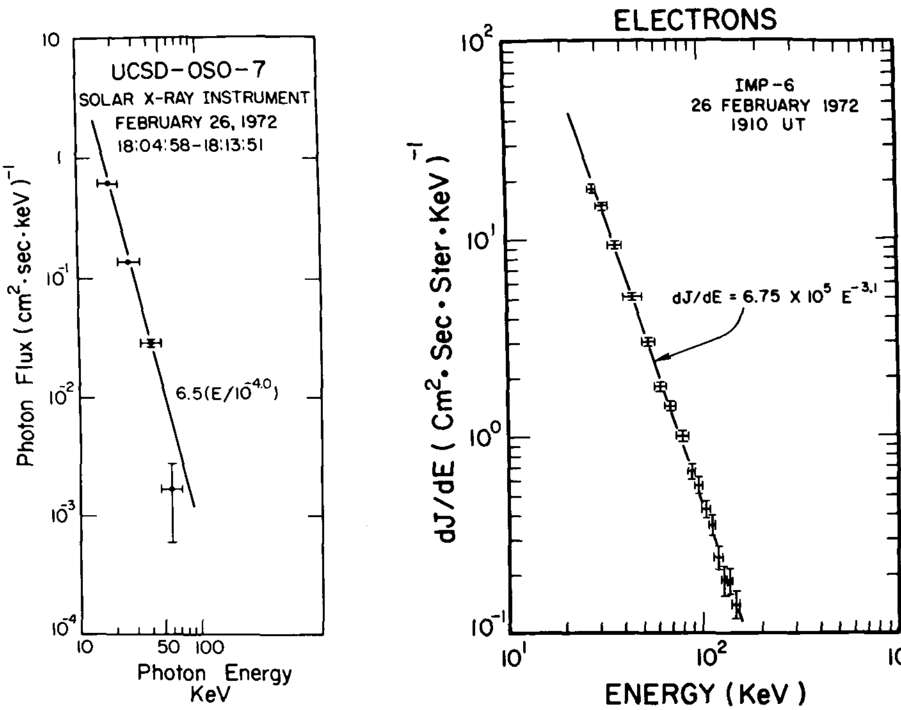


Fig. 4. The spectra of hard X-rays and electrons observed at 1 AU for the same flare event. The photons fit a power law spectrum $dJ(h\nu)/d(h\nu) = A(h\nu)^{-\gamma}$ where $\gamma = 4.0 \pm 0.3$, while the electrons fit a spectrum $dJ/dE = 6.75 \times 10^5\ E^{-3.1}$. Since $dn/dE = v\ dJ/dE$ where v is the electron velocity, the electron fit a *density* spectrum $dn/dE \propto E^{-\delta}$ with $\delta = 3.6 \pm 0.1$. These two spectra are consistent with thin target emission under the assumption the escaping electrons have the same spectrum as the accelerated electrons.

(see Lin and Hudson, 1971; Kane and Lin, 1972) with a thin target model. The thin target case is also consistent with the location of the acceleration region ($n_i \lesssim 10^{10}$ cm^{-3}) derived from considerations of the low energy electron spectrum observed at 1 AU.

In favor of thick target processes we note that if non-relativistic electrons penetrate to the dense ($n_i \gtrsim 10^{12}$ cm^{-3}) regions of the chromosphere-corona boundary and below, they could produce the observed EUV and perhaps provide the energy for heating the Hα flare region through collisional loss (and possibly even heat the white light flare region) (Hudson, 1972). The close time coincidence between the hard X-ray spike and the EUV spike (Kane and Donnelly, 1971) is consistent with such an interpretation. At those densities the thick-target approximation would certainly be appropriate.

There are several possible ways of reconciling the observations in support of thick and thin target. One possibility is that the electrons injected into the interplanetary medium may have a spectrum modified from that of the electrons initially accelerated. From comparisons of the total energy in accelerated electrons derived from the X-ray observations and the total energy in escaping electrons, Lin and Hudson (1971) found that the escape efficiency is only ~ 0.1 to 1%. Thus the probability of escape of the electrons into the interplanetary medium may be a function of electron energy. We do not favor this possibility in view of (a) the relatively small amounts of matter which are traversed by the escaping electrons, and (b) the otherwise coincidental agreement between the X-ray and electron spectra.

A second possibility is that electrons of low energies, say below ~ 10 keV, are described by the thick target approximation while higher energy electrons are in an essentially thin target situation (Kane, 1973). This dichotomy could arise, for example, if the electrons are accelerated and contained by a magnetic 'bottle' in a low density, $n_i \lesssim 10^{10}$ cm^{-3}, region. Electrons only appear in high density, $n_i \gtrsim 10^{10}$ cm^{-3}, regions near the feet of the magnetic bottle if they are scattered into the loss cone. Since the amount of scattering is a strongly decreasing function of energy, essentially only the low energy electrons will be dumped into the loss cone. This interpretation is consistent with the observations which show that the correspondence between rising portion of the EUV emission and the rising portion of the non-thermal X-rays is best for the lowest energy, ~ 10 keV, X-rays.

A third possibility is that the acceleration of the escaping electrons is separate from the acceleration of the electron producing X-rays. However the time of injection of the electrons into the interplanetary medium, which can be obtained accurately ($\lesssim 10$ min) by analyses of the velocity dispersion observed during the onset of electron events at 1 AU, is clearly between the Hα onset and maximum, i.e., at the time of the flash phase X-ray and radio flare phenomena. The duration of the injection is inferred to be $\lesssim 3$ min from the duration of the highly scatter-free events. Furthermore, escaping electrons generate type III emission. Type III emission occurs exactly at the time of X-ray bursts to within seconds when both are observed from flares (Kane, 1972).

These considerations indicate that the >20 keV electron spectrum observed at 1 AU is probably an essentially undistorted sample of the electrons accelerated in the flare.

5. Two Stage Acceleration

Two types of electron spectra are observed (Lin, 1970). For events which are unaccompanied by energetic protons* and relativistic electrons, i.e., pure non-relativistic electron events, the electron spectrum can be fit to a power law with exponent from ~ 2 to 5, usually with a steepening to >5 at ~ 100–200 keV (Figure 5). Events which are accompanied by energetic protons usually have electron spectra which extend

* Above a threshold of ~ 0.3 (cm^2 s ster)$^{-1}$ above 10 MeV.

smoothly in a power law to relativistic energies (Figure 6). These two types of spectra, one with a 'cut-off' and one without, suggest two stages of acceleration, one a flash phase acceleration of mainly just ~5–100 keV electrons, and the second an acceleration of protons and electrons to high, even relativistic energies which occurs only in some flares. This concept is further supported by spacecraft observations which in-

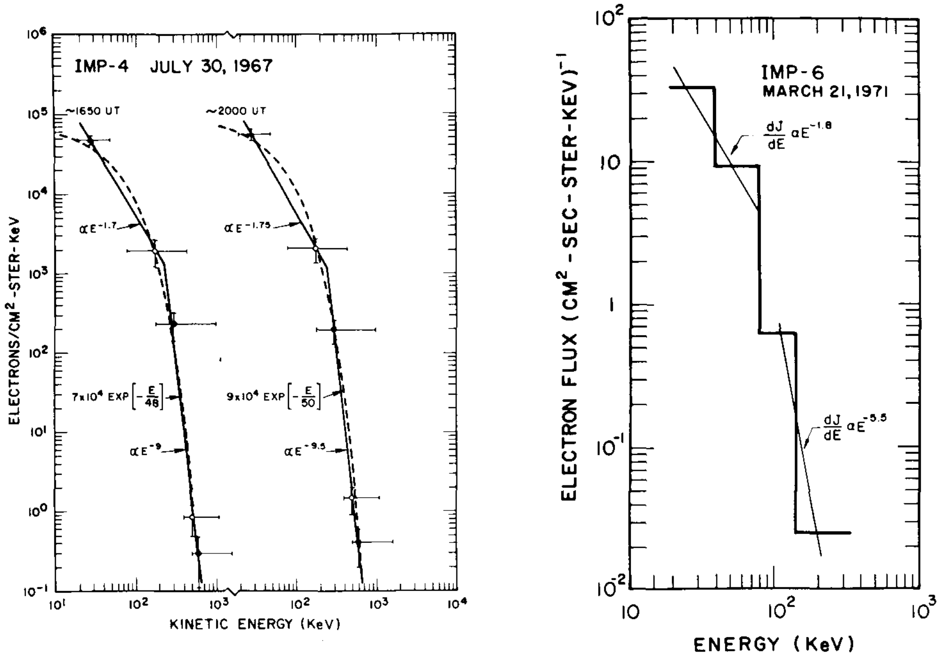


Fig. 5. The electron energy spectra of the events of Figure 2 (from Wang *et al.*, 1971) illustrating the steepening in the spectrum above ~100–200 keV typical of pure electron events. Note the wide variation in flux from the 30 July 1967 events to the 21 March 1971 event.

dicate that relativistic electrons and energetic protons are usually injected into the interplanetary medium ≳10 min *after* the non-relativistic electrons (Figure 7) (Sullivan, 1974; Simnett, 1974; Lin and Anderson, 1967).

Two stages are sometimes observed in the hard X-ray event accompanying energetic proton flares (Figure 8) (Frost and Dennis, 1971). The X-ray energy spectrum shows a cut-off at ~100 keV for the flash phase but no energy cut-off even to the limits of their observation (≳250 keV) in the long second phase. That phase starts at the onset of the type II burst. The radio, X-ray, and particle observations are generally consistent with the acceleration of particles in the second state by the type II shock front in the *corona* by a stochastic Fermi-type mechanism. The X-ray event of 30 March 1969 (Figure 8) was associated with a behind-the-limb flare (~W 110°) so the origin of the X-ray burst must be in the corona. Interplanetary shocks and the Earth's bow shock both accelerate particles, electrons and nuclei, up to energies of ~10^2 keV (Fan *et al.*, 1964; Anderson, 1965; McGuire *et al.*, 1972) and ~10 MeV (Palmeira *et al.*, 1971) re-

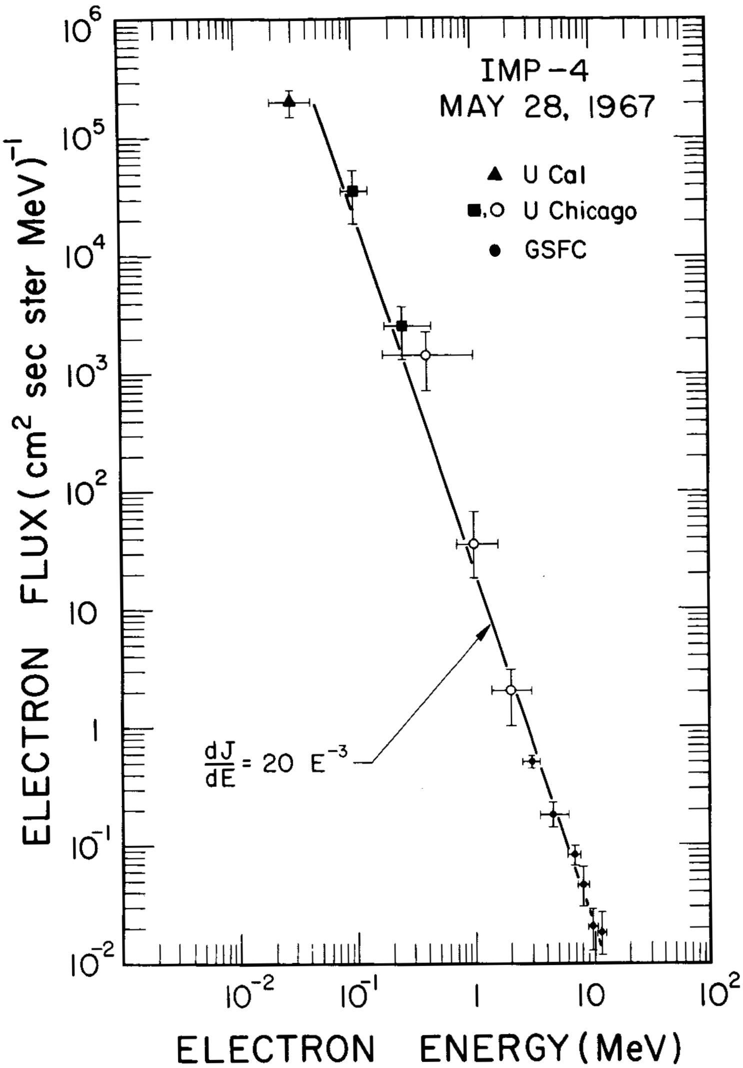


Fig. 6. The differential electron energy spectrum for the mixed electronproton event of 28 May 1967, compiled from four different detector systems aboard IMP-4. The lowest energy point is obtained from Geiger Müller detector observations (Lin, 1970). The University of Chicago points are from solid-state detector telescopes (Sullivan, 1973), and the points above ~2 MeV are from the Goddard Space Flight Center range and energy loss scintillation detectors (Simnett, 1971). The points fit to a single power over three decades in energy, even though they are from several different detectors with different view directions.

spectively. In the much higher magnetic fields and densities near the Sun it seems likely that substantially higher particle energies will be attained.

6. Generation of Type III Radio Emission

It is well established that energetic electrons generate most (if not all) of the non-thermal radio emission observed from the Sun. Detailed calculations of gyro-synchrotron and synchrotron emission of energetic electrons in solar magnetic fields give generally good agreement with observations of impulsive microwave bursts and type IV emission. However until recently the various theoretical treatment of type III emission differed as to the exciter, whether electrons or protons or waves, and numbers

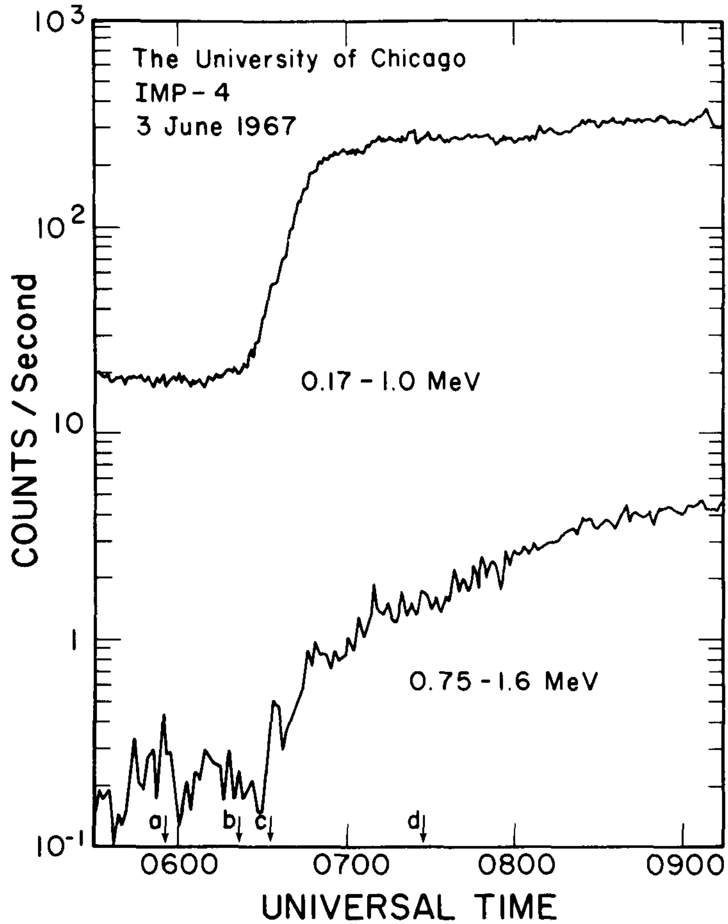


Fig. 7. The delay of the relativistic (0.75–1.6 MeV) electrons vs the non-relativistic electrons (0.17–1 MeV) is illustrated here. The onset of the soft 2–12 Å flare X-rays is denoted by (a), the 0.17–1.0 MeV electron onset by (b), the 0.75–1.6 MeV onset by (c), and the 9.6–18 MeV proton onset by (d) (taken from Sullivan, 1973).

of particles needed varied over 10–12 orders of magnitude (Evans *et al.*, 1971). Thus direct observations of the exciter, and simultaneously the type III emission generated by them at 1 AU, are of critical importance in establishing a firm theoretical base for the emission processs.

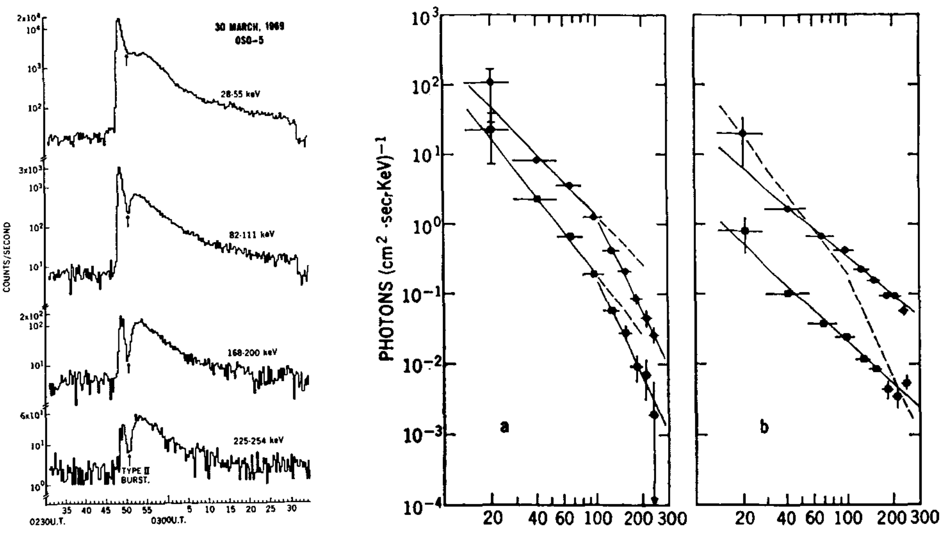


Fig. 8. The two-stage hard X-ray burst of 30 March 1969 (from Frost and Dennis, 1971), which was followed by an intense mixed electron-proton event observed at 1 AU. This X-ray event presents clear evidence for two-stage acceleration. Photon energy spectra during the initial X-ray burst (shown in the middle panel marked 'a') have a spectrum steepening above ~100 keV, fairly typical of flash phase events. The spectra during the second phase (shown in the right panel marked 'b') which starts at the time of intense type II emission (Smerd, 1970) shows a smooth and very hard spectrum to the upper limits of the X-ray detector's energy range (~300 keV).

6.1. Physical Mechanism

Type III solar radio bursts are the most common type of impulsive phenomena observed from the Sun. These bursts are characterized by a rapid frequency drift from high to low frequencies, and occasionally by the presence of two bands of emission, one at approximately twice the frequency of the other (see reviews by Wild, *et al.*, 1963; Wild and Smerd, 1972). A theoretical basis for the plasma hypothesis for type III solar radio bursts (Wild, 1950) was first introduced by Ginzburg and Zheleznyakov (1958), and although it has been developed and refined in the intervening years, the basic ideas have remained unchanged (see review by Smith, 1973). A group of fast particles injected near the Sun generate longitudinal electron plasma waves at frequencies near the local plasma frequency as they pass through the coronal plasma. These plasma waves then scatter off ion density fluctuations to produce electromagnetic radiation near the plasma frequency (fundamental), and off other plasma waves to produce emission at twice the plasma frequency (2nd harmonic). As the fast particles go upward in the corona and into the interplanetary medium the radio emission will drift from high to low frequencies. Typical drift rates for these bursts indicate velocities of ~0.3–0.5 c for the particles where c is the speed of light.

The plasma waves are generated through a coherent Cerenkov plasma process. In order to produce plasma waves more rapidly than they are damped, the velocity distribution of the fast particles must have a positive slope, that is, a peak must exist in the non-thermal particle velocity distribution. Observations of the characteristics of type III bursts and the particles that excite them will thus provide a test of basic beamplasma and mode-mode coupling theory over a wide range of plasma conditions.

6.2. THE TYPE III BURST EXCITER

Wild *et al.* (1954) were first to suggest that energetic protons might be the exciters of type III emission. More recently Smith (1970) summarized the theoretical difficulties of stabilizing a spatially unbounded and homogeneous electron stream, and noted that a proton stream can be stabilized. Smith (1970) suggested that the $\sim$20–100 MeV protons which might produce the burst at the Sun were so few in number that after diffusion in the interplanetary medium their fluxes would be too low to observe at 1 AU. However, the observations of type III bursts generated near 1 AU imply that substantial fluxes of protons sufficient to produce emission should be observed *at 1 AU*. Such fluxes are not generally observed except in large proton events which are quite rare compared to electron events.

The theoretical difficulties for electron streams can apparently be overcome by considering a spatially bounded stream with inhomogenities in the front and back (Zaitsev *et al.*, 1972). In addition direct observations of the electron velocity distributions at 1 AU show that peaked distributions do exist, contrary to the theoretical predictions otherwise.

A very highly significant correlation, almost one to one, exists between intense kilometric wavelength type III's and >20 keV electrons observed at 1 AU from flares located in the western solar hemisphere (Alvarez *et al.*, 1972). In addition storms of weak type III bursts are observed at hectometric wavelengths which appear to be closely related to type I storms at metric wavelengths (Fainberg and Stone, 1970). These type III storms are accompanied by non-impulsive co-rotating $\gtrsim 20$ keV electron fluxes observed at 1 AU.

Recently Frank and Gurnett (1972) and Lin *et al.* (1973a) have reported observations at 1 AU of energetic electrons and type III burst emission at the low frequencies characteristic of the near 1 AU plasma environment. Frank and Gurnett (Figure 9) did not observe radio emission simultaneously with the arrival of the 5–6 keV electrons, which, in their interpretation, are the exciters of the emission. Rather they assumed that the radiation is generated primarily at the fundamental, i.e., local plasma frequency, and that the lowest frequency radiation they observe, $\sim$31 kHz, originates some distance away from 1 AU. Approximately 2600 s after the onset of the 31 kHz emission the $\sim$6 keV electrons are observed to arrive. Although the fundamental emission generated at 1 AU ($\sim$20 kHz) is not observed, the authors note that the calculated time of onset of the fundamental emission is in agreement with the arrival of the $\sim$6 keV electrons ($v \approx 0.15\ c$). Since the drift rates of bursts near the Sun indicate velocities of $\sim$0.3 c, they conclude that deceleration of the electrons may be substantial.

Lin *et al.* (1973a) located the position of the type III burst emission at each fre-

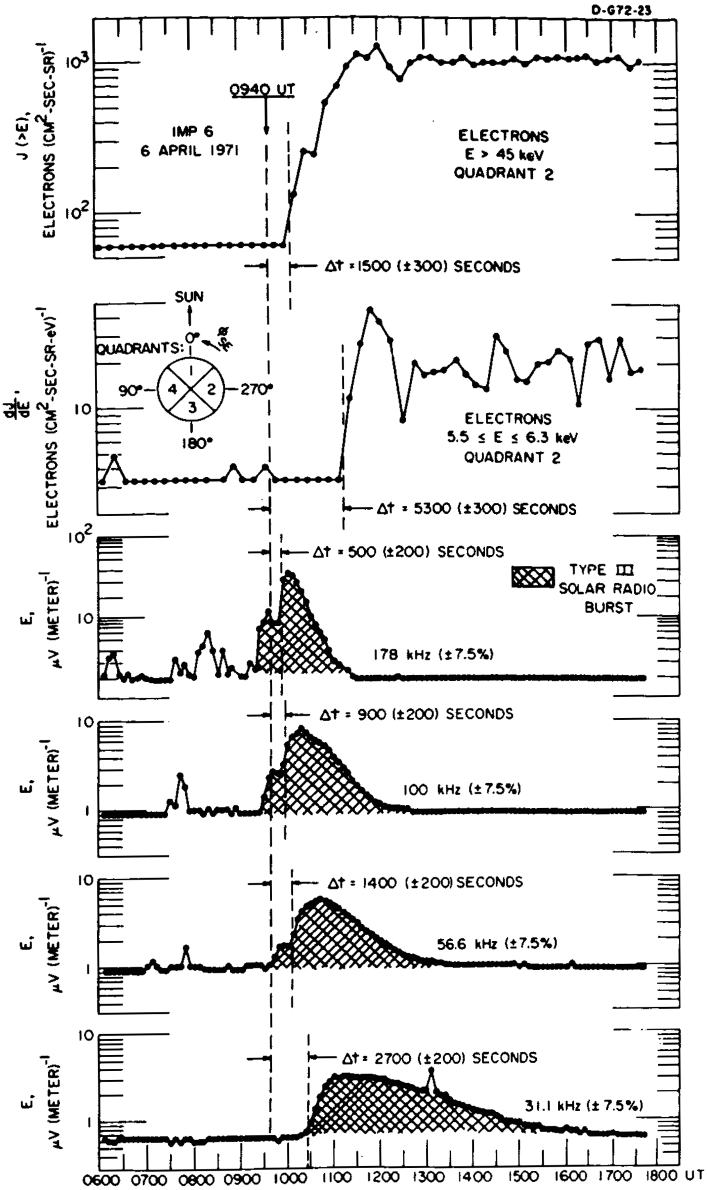


Fig. 9. Simultaneous electron and type III burst observations at 1 AU (from Frank and Gurnett, 1972). The local electron plasma frequency is ~20 kHz, but no oscillations were observed at that frequency. If the radio emission is at the fundamental of the plasma frequency then the 5.5–6.3 keV electrons would be likely candidates for the burst exciter. If, on the other hand, the radio emission is at the second harmonic of the plasma frequency then higher energy, $\gtrsim 10$ keV, electrons would coincide with the emission (see text).

quency (Figure 10) from the spin modulation of the radio signal. The emission originating at 1 AU was then compared with the particle data. Their results show (Figure 11) that the onset of the emission located at 1 AU corresponds to the arrival of electrons of ~100 keV energy. The subsequent build up of the radiation corresponds to the arrival of lower energy electrons, until maximum is reached when ~10 keV

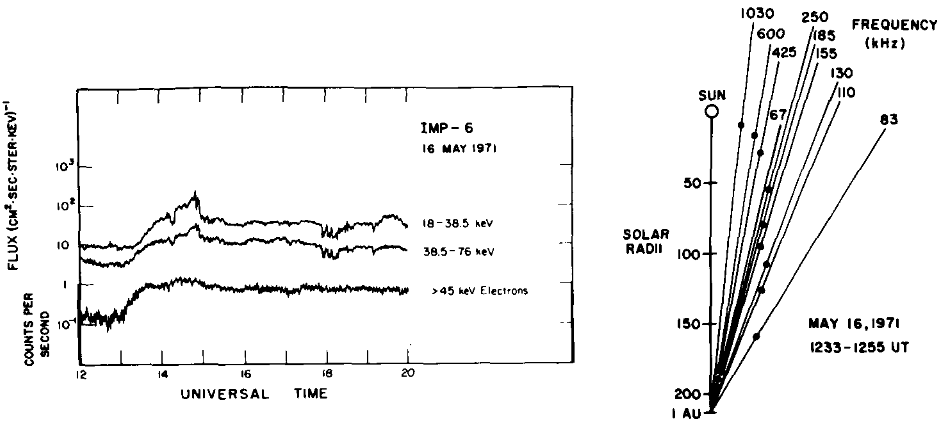


Fig. 10. (a) The May 16, 1971 event. The upper two channels are sensitive to low energy protons as well as electrons while the lower channel is sensitive only to electrons. Some upstream terrestrial protons are observed ~1430–1500 UT, a time well after the period of analysis. (b) The trajectory in the interplanetary medium of the type III burst of 16 May 1971, determined from the spin modulation of the observed radio signal at different frequencies.

electrons first arrive. By the time 6 keV electrons arrive the emission is already decaying rapidly.

Since the propagation of these electrons may differ markedly from event to event, depending on the changeable scattering characteristics of the interplanetary medium, two events were compared. One event was scatter-free, the other diffusive. The frequency drift rate of the interplanetary type III bursts was more rapid for the scatter-free event and less rapid for the diffusive event, corresponding nicely to the difference in the computed distance traveled for the first arriving electrons of 1.4 AU in the scatter-free event and 1.7 AU for the diffusive event.

The evidence from radio studies indicate that the second harmonic emission rather than fundamental is predominant for low frequency type III radiation (Fainberg *et al.*, 1972; Smith, 1972; Malitson *et al.*, 1973; Haddock and Alvarez, 1973). Thus, the frequencies of the near 1 AU radiation are ~twice the local plasma frequency. Under the assumption of second harmonic radiation the observations of Frank and Gurnett would be in close agreement with Lin *et al.* (1973a).

Even under the second harmonic hypothesis some *apparent* deceleration is observed (Fainberg *et al.*, 1972). This can be attributed to scattering of the electrons as they propagate outward in the interplanetary medium. Such scattering will lower the apparent velocity along a smooth spiral field line.

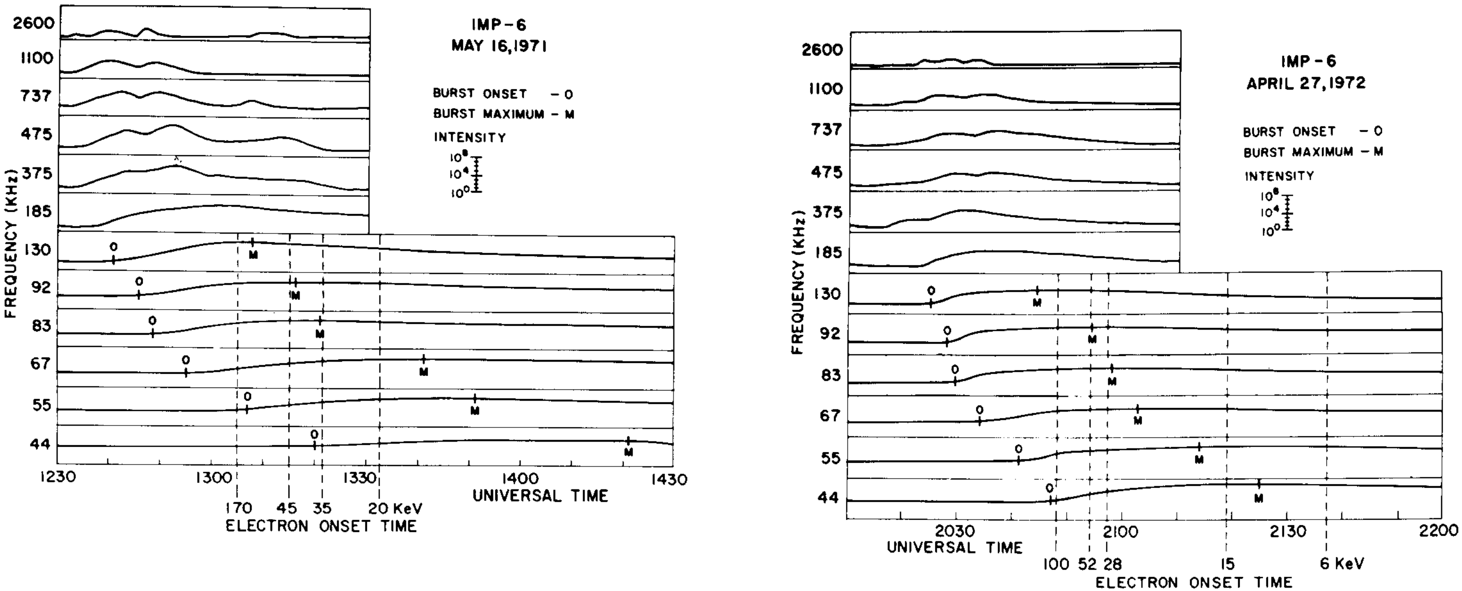


Fig. 11. (a) The type III burst of 16 May 1971 observed at frequencies from 2.6 MHz to 44 kHz. The 55 kHz emission originates closest to 1 AU (see Figure 10). (b) The type III burst of 27 April 1972. Here the 44 kHz emission is closest to 1 AU.

6.3. Comparison with Theories

Using calculated incoherent Cerenkov emission efficiency (Cohen, 1958) and a maximum estimated value for the efficiency of coupling to EM radiation, we find that the emission from incoherent Cerenkov processes calculated for the observed electron fluxes is insufficient to account for the observed emission.

The peaked distribution needed to generate coherent Cerenkov radiation is clearly present in the electrons (Figure 12). This peaked distribution arises from velocity

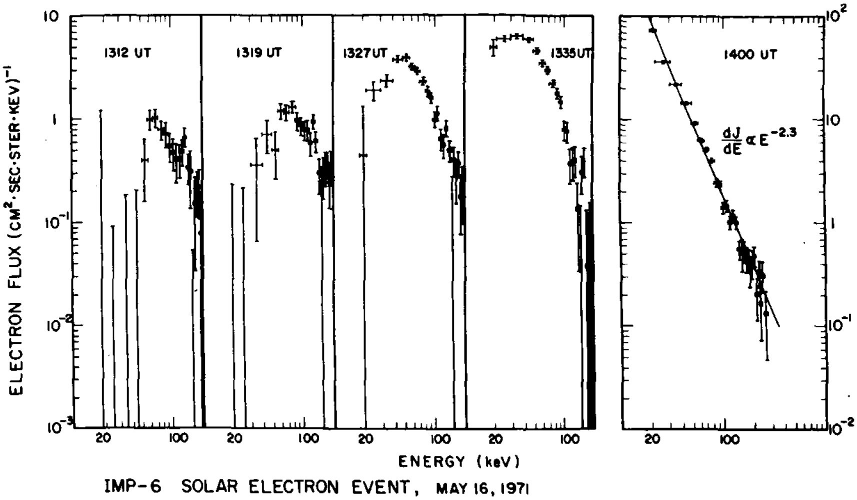


Fig. 12. The energy spectrum of the electrons in the 16 May 1971 event at different times during the onset.

dispersion rather than an initially peaked distribution of injected electrons at the Sun, since the injection spectrum at the Sun is observed to extend down to $\lesssim 5$ keV without a peak.

The radio emission is observed to increase in intensity until the peak in the electron distribution falls below ~ 15 keV. The radio burst begins a constant exponential decay, independent of the behavior of the electron flux, after the peak in the electron distribution goes below ~ 10 keV.

It should be noted that the coherent Celenkov processes, at least in these observations, are not so strong as to substantially modify the peaked distribution. That distribution evolves essentially as would be anticipated from velocity dispersion alone. The implication is that the time scale for relaxation of the beam through the wave-particle interaction is long compared to the time scale for evolution of the distribution by velocity dispersion. This is contrary to the expectation of most theories for type III emission.

One factor not considered in most theoretical treatments which may be relevant in explaining the long time scale for relaxation is the angular distribution of the electrons. The average distribution for the 16 May 1971 event during the period of generation of type III emission is shown in Figure 13. The distribution is not just outward along the field line as assumed by most theoretical models, but instead is relatively isotropic, with a maximum to minimum ratio of ~2:1. This angular distribution presumably arises from the pitch-angle scattering of the electrons by irregularities in the interplanetary magnetic field.

Table III lists some characteristics of the low frequency type III burst at 1 AU

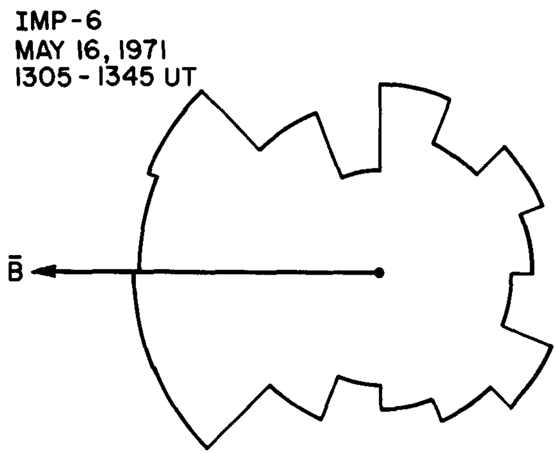


Fig. 13. The pitch angle distribution of the >45 keV electrons during the onset of the 16 May 1971 event. The approximate direction of B is indicated by the arrow.

TABLE III

Relevant parameters of type III bursts and burst exciters at 1 AU

(1) Type III burst	
(a) Typical frequencies of observation (2nd harmonic)	~40–60 kHz
(b) Intensity of emission	$\sim 3 \times 10^{-17}$ W m^{-2} Hz
(c) Cross-sectional area of source	$\sim 10^{26}$ cm^2
(d) Average emissivity of source	$\sim 6 \times 10^{-24}$ erg cm^{-3} s^{-1}
(2) Fast electron exciters	
(a) Typical energy of electrons	~10–100 keV
(b) Density of fast electron at burst maximum	$\sim 10^{-6}$ cm^{-3}
(c) Total number of fast electrons producing emission	$\sim 10^{33}$
(d) Average rate of energy loss to electromagnetic emission (per electron)	$\sim 4 \times 10^{-9}$ keV s^{-1}
(e) Average interparticle distance	$\sim 10^2$ cm
(3) Ambient plasma medium (solar wind)	
(a) Mean densities	5–10 cm^{-3}
(b) Mean temperatures – electron	$\sim 1.2 \times 10^5$ K
– proton	$\sim 7 \times 10^4$ K
(c) Debye length	$\sim 10^3$ cm

and some parameters of the emission process. It is clear from these preliminary studies that a quantitative plasma theory for type III bursts can be obtained through further observations of this kind. Such observations and theory, besides illuminating problems of plasma physics otherwise impossible to study, will also be applicable for the quantitative interpretation not only of solar flare radio phenomena, but possibly of galactic and extragalactic radio emission as well.

Acknowledgements

I wish to acknowledge discussions with Drs S. R. Kane and K. A. Anderson. This research was funded in part by NASA grant NGL-05-003-017.

References

Alvarez, H., Haddock, F., and Lin, R. P.: 1972, *Solar Phys.* **26**, 468.
Anderson, K. A.: 1965, *Proc. 9th International Conf. on Cosmic Rays*, London, p. 520.
Brown, J. C.: 1971, *Solar Phys.* **18**, 450.
Cohen, M. H.: 1958, *Phys. Rev.* **123**, 711.
Datlow, D.: 1971, *Solar Phys.* **17**, 436.
Datlowe, D. and Lin, R. P.: 1973, *Solar Phys.* **32**, 459.
Evans, L. G., Fainberg, J., and Stone, R. G.: 1971, *Solar Phys.* **21**, 1968.
Fainberg, J., Evans, L. G., and Stone, R. G.: 1972, *Science* **178**, 743.
Fainberg, J. and Stone, R. G.: 1971, *Astrophys. J.* **164**, L123.
Fan, C. Y., Gloeckler, G., and Simpson, J. A.: 1964, *Phys. Rev. Letters* **13**, 149.
Frank, L. A. and Gurnett, D. A.: 1972, *Solar Phys.* **27**, 446.
Frost, K. J. and Dennis, B. R.: 1971, *Astrophys. J.* **165**, 655.
Ginzburg, V. L. and Zheleznyakov, V. V.: 1958, *Soviet Astron.* **AJ2**, 653.
Haddock, F. T. and Alvarez, H.: 1973, *Solar Phys.* **29**, 183.
Hudson, H. S.: 1972, *Solar Phys.* **24**, 414.
Kane, S. R.: 1972, *Solar Phys.* **27**, 174.
Kane, S. R.: 1973, *Bull. Am. Astron. Soc.* **5**, 274.
Kane, S. R. and Anderson, K. A.: 1970, *Astrophys. J.* **162**, 1003.
Kane, S. R. and Donnelly, R. F.: 1971, *Astrophys. J.* **164**, 151.
Kane, S. R. and Lin, R. P.: 1972, *Solar Phys.* **23**, 457.
Krimigis, S. M.: 1965, *J. Geophys. Res.* **70**, 2943.
Lin, R. P.: 1970a, *J. Geophys. Res.* **75**, 2583.
Lin, R. P.: 1970b, *Solar Phys.* **12**, 209.
Lin, R. P.: 1971, *Conf. Papers, 12th International Conf. on Cosmic Rays*, Hobart, Tasmania, Australia, **5** 1805.
Lin, R. P.: 1973, in R. Ramaty and R. G. Stone (eds.), *Proc. Symposium on High Energy Phenomena on the Sun*, Goddard Space Flight Center, Greenbelt, Maryland, September 1972, p. 439.
Lin, R. P. and Anderson, K. A.: 1967, *Solar Phys.* **1**, 446.
Lin, R. P. and Hudson, H. S.: 1971, *Solar Phys.* **17**, 412.
Lin, R. P., Evans, L. G., and Fainberg, J.: 1973a, *Astrophys. Letters* **14**, 191.
Lin, R. P., Wang, J. R., and Fisk, L. A.: 1974, *J. Geophys. Res.*, in preparation.
Malitson, H. H., Fainberg, J., and Stone, R. G.: 1973, *Astrophys. Letters* **14**, 111.
McGuire, R. E., Anderson, K. A., Chase, L. M., Lin, R. P., and McCoy, J. E.: 1972, *Trans. AGU* **53**, 1086
Palmeira, R. A. R., Allum, F. R., and Rao, U. R.: 1971, *Solar Phys.* **21**, 204.
Parker, E. N.: 1963, *Interplanetary Dynamical Processes*, Interscience Publishers, Chapter 8.
Simnett, G. M.: 1971, *Solar Phys.* **20**, 448.
Simnett, G. M.: 1974, *Space Sci. Rev.* **16**, 257.
Smerd, S. F.: 1970, *Proc. Astron. Soc. Australia* **1**, 305.
Smith, D. F.: 1970, *Solar Phys.* **15**, 202.

Smith, D. F.: 1972, in R. Ramaty and R. G. Stone (eds.), *Proc. Symposium on High Energy Phenomena on the Sun*, Goddard Space Flight Center, Greenbelt, Maryland, September, 1972, p. 558.
Sullivan, J. D.: 1974, *J. Geophys. Res.*, to be published.
Syrovatskii, S. I. and Shmeleva, O. P.: 1972, *Soviet Astron. AJ* **16**, 273.
Trubnikov, B. A.: 1965, *Rev. Plasma Phys.* **1**, 105.
Wang, J. R., Fisk, L. A., and Lin, R. P.: 1971, *Conf. Papers, 12th International Conference on Cosmic Rays*, Hobart, Tasmania, Australia, **2**, 438.
Wild, J. P.: 1950, *Australian J. Sci. Res.* **A2**, 541.
Wild, J. P., Murray, J. D., and Rowe, W. C.: 1954, *Australian J. Phys.* **7**, 439.
Wild, J. P., Smerd, S. F., and Weiss, A. A.: 1963, *Ann. Rev. Astron. Astrophys.* **1**, 291.
Wild, J. P. and Smerd, S. F.: 1972, *Ann. Rev. Astron. Astrophys.* **10**, 159.
Zaitsev, V. V., Mityakov, N. A., and Rapoport, V. O.: 1972, *Solar Phys.* **24**, 444.

NGL-05-003-017 D2

Mapping of Lunar Surface Remanent Magnetic Fields by Electron Scattering

R. P. Lin, R. E. McGuire*, H. C. Howe†, K. A. Anderson*

Space Sciences Laboratory
University of California
Berkeley, California 94720

and

J. E. McCoy

Lyndon B. Johnson Space Center
Houston, Texas 77058

We present preliminary results of a new technique to detect and measure weak and small scale lunar magnetic fields (Howe et al.).[1] This technique makes use of the fact that charged particles are easily mirrored or scattered in moving to a region of increased magnetic field strength. The particles, in effect, provide a probe along the external magnetic field line down to the lunar surface. This probe is sensitive to increases in field strength along its path. If no surface magnetization is present, the particles are guided by the external magnetic field into the lunar surface where they are absorbed (except for a few percent which are Coulomb backscattered from surface material). If surface remanent magnetism is present, the total field strength increases as the particles

*Also Physics Department.

†Present address: Thermonuclear Division, Oak Ridge National Laboratory Oak Ridge, Tennessee 37830

enter the region of remanent magnetism, causing a fraction of the particles to mirror or backscatter with an intensity that increases with the average strength of the total surface field.

This effect has repeatedly been observed for electrons by the Apollo 15 and 16 Particles and Fields Subsatellites (PFS 1 & 2). When the reflections are strong the on-board magnetometer records a distinct magnetic effect. All the effects we attribute to electron mirroring and scattering are fixed with respect to lunar surface features. The spatial resolution of this technique is limited only by the gyro-radius of the particles used as probes. Electrons, because of their small gyro-radii, are ideal for this application.

The backscattering of the electrons depends on the scale size of the region of remanent magnetism. Monte Carlo simulations using direct force equation trajectory calculations have been performed to estimate the scattering from various magnetic configurations. The reflection coefficient for fluxes in the 45° -90° pitch angle sector is found to be well approximated by the assumption of adiabatic particle motion under the average surface field (averaged over the electrons gyro orbit). This assumption provides a simple way of estimating the total field strength at the lunar surface:

$$B_{surf} = B_0/\sin^2 \alpha_c$$

where α_c, the limiting pitch angle to be backscattered is calculated from the observed ratio of fluxes R between the incident and returning 45° -90° pitch angle fluxes

$$R = \frac{\int_{\alpha_c}^{90^\circ} j(\alpha)\ \sin\alpha\ d\alpha}{\int_{45^\circ}^{90^\circ} j(\alpha)\ \sin\alpha\ d\alpha}$$

where B_0 is the ambient magnetic field, α is the electron pitch angle, $j(\alpha)$ is the directional electron flux, $B_{surf} = |\overline{B}_0 + \overline{B}_{rem}|$.

Using the ratio of backscattered to incoming electron flux at 14 keV we have mapped the remanent fields (B_{rem}) near the PFS-2 ground track with a spatial resolution of $\sim$ 40 km and a sensitivity of 0.1 γ average field over the 40 km x 40 km resolution element. The results are shown in the form of color overlay maps (Figure 1) for the two cases of the external field pointing into and out of the lunar surface (B_{in} and B_{out}) respectively.

The 14 keV electron maps confirm the generally stronger fields reported in the lunar highlands reported by Sharp et al.[2] and also show evidence of sharp changes in the lunar remanent magnetism in less than one resolution element (40 km).[3]

During real time tracking of the PFS, the $\frac{1}{2}$ keV electron flux is read out five times during the satellite spin period of five seconds. These high temporal resolution, low energy data allow mapping of the surface remanent fields to a spatial resolution of better than 10 km. Some preliminary results of the ½ keV mapping have been presented elsewhere.[4] This technique can be extended to fields as weak as 10^{-7} Gauss at the lunar surface and to spatial resolution of $\leqslant$ 3 km.

Acknowledgments

We thank P. J. Coleman and C. T. Russell for supplying the magnetometer data. This research was supported in part by NASA Contract NGS-7127 and Grant NGL-05-003-017.

References

1. Howe, H. C., R. P. Lin, R. E. McGuire and K. A. Anderson, Energetic electron scattering from the lunar remanent magnetic field, Geophys. Res. Lett., 1, 101, 1974.
2. Sharp, L. R., P. J. Coleman, B. R. Lichtenstein and C. T. Russell, Orbital mapping of the lunar magnetic field, The Moon, 7, 322, 1973.
3. McCoy, J. E., K. A. Anderson, R. P. Lin, H. C. Howe and R. E. McGuire, Lunar remanent magnetic field mapping from orbital observations of mirrored electrons, Lunar Interactions Conference (Abstract), p. 52, 1974.
4. Anderson, K. A., R. P. Lin, R. E. McGuire, H. C. Howe and J. E. McCoy, High resolution mapping of the lunar surface remanent magnetic field by energetic electron scattering, Lunar Science VI, Abstracts of the Sixth Lunar Science Conference, p. 12, 1975.

Figure Caption

Figure 1. This figure shows lunar surface magnetic field strengths obtained from the electron scattering technique (see text, page 3). The field strengths are average values over a 40 km x 40 km region, and are plotted from $< 0.1\gamma$ ($1\gamma = 10^{-5}$ Gauss) to $> 1.6\gamma$ increasing by factors of two. The data shown are from magnetotail and magnetosheath passes of the Apollo 16 Subsatellite. Separate maps have been constructed for times when the external magnetic fields are directed outward and inward with respect to the lunar surface.

THE MATURATION OF LUNAR DUST GRAINS. J.L. Bertaux°, J.P. Bibring*, J. Borg*, A.L. Burlingame+, Y. Langevin*, M. Maurette*, F.C. Walls+. °Laboratoire d'Aeronomie du CNRS, 91 Verrières-le-buisson. *Laboratoire René Bernas, 91406 Orsay. +Space Science Laboratory, UC Berkeley, 94720 California.

I. INTRODUCTION. As a result of lunar "weathering" processes active at the interface between the moon and its space environment, "maturation" proceeds both in modifying the surface physicochemical properties of lunar dust grains and in injecting new grains in the regolith. In this report we first present a "dynamical" history of maturation obtained by "theoretically" generating the individual motions of dust grains in the top layers of the lunar regolith. Then we use results of simulation experiments to define solar wind maturation as well as to propose a new sampling technique for extracting grains in lunar core tubes which should greatly help in deciphering the past activity of this type of maturation. Finally we try to correlate several bulk properties of the lunar regolith to specific "weathering" processes active during maturation.

II. THE SIMULATED "DYNAMICAL" HISTORY OF MATURATION.

In our Solmix computations (1) intended to predict the most probable motion of lunar dust grains in the regolith, we define 2 parameters, Δ_{LS} and N_S, which are the key "dynamical" factors involved during maturation.

II.1. Δ_{LS}, the lunar "skin". Let us consider a grain with a radius, r, at an initial depth, d_o, in the top strata of the regolith. As a result of the gardening due to the meteorite rainfall such a grain has a probability, $S(r, d_o, t_o = \overline{\lambda}_b, n_s)$ to get exposed more than n_s times on the top surface of the regolith, during a time interval, $\overline{\lambda}_b = 25.10^6$ years, which represents the meanlife of the top strata against burial by another strata (see chapter III.1, in ref.1). From the sharp drop of $S(n_s)$ when d_o increases we infer the following conclusions: 1. when $d_o \lesssim 5$mm the 50 μ-grains ($r \sim 50\mu$) have a high probability ($\gtrsim 80\%$) of being exposed in more than 5 different orientations ($n_s > 5$) at depth 0 ; 2. $S(n_s)$ decreases with the size of the grains and only 50 % of the 1 μ-grains within the same surface layer ($d_o \lesssim 5$mm) have been exposed at least once at depth 0, during the $\overline{\lambda}_b$ time interval. However as soon as the 1 μ-grains ($r \sim 1\mu$) get exposed at depth 0 they suffer additional changes in orientation ($\simeq 0.1$ per year) due to the momentum imparted by the lunar winds ; 3. therefore we will define this most superficial layer of the top strata with a thickness $\sim$ 5mm, as the lunar "skin", Δ_{LS}, in which maturation will proceed as a high proportion of grains, at depths $d_o \lesssim 5$mm, get homogeneously exposed on "all their faces" to lunar weathering processes.

II.2. N_S, the number of "progenitor" top strata, through which the grains have been cycled before their most recent deposition in the regolith - for strata with a thickness $\Delta_S \gtrsim 1$cm, N_S has a value of about $2.5/10^9$ years (ref.1, Chapter III.1). This dynamical factor is also involved during maturation for the following reasons : i. the necessary condition for a "fresh" 50 μ-grain just admixed at any depth into the top strata to proceed to maturation is to get incorporated within Δ_{LS}, and the probability of this event is $\simeq 0.8.\Delta_{LS}/\Delta_S$. For an average value $\Delta_S \simeq 3$cm (ref.1, chapter III.1) such a grain should then be successively cycled through about 7 "progenitor" strata, to be incorporated within a lunar "skin" ; ii. during the N_S cycling, glassy grains are certainly injected in the regolith and accretionary particles could probably get stuck to the external surface of the coarser grains.

III. MATURITY INDICES

It is difficult to define a meaningful experimental index of maturity. Indeed the "bulk" maturation of the lunar soil results from the combination of a variety of lunar "weathering" effects (implanted species and radiation damage coatings due to the solar wind ; solar flare tracks ; accretionary particles ; glassy agglutinates ; impact craters in the sub-micron range) which correspond to different processes. Furthermore the extent of maturation related to each one of these processes is likely to depend on both the size and the nature of the grains that are used to measure the index of maturity.

This last difficulty is particularly well illustrated for solar wind maturation which only occurs for grains directly exposed at depth 0. In fact the active depth associated with this type of maturation is given by the thickness of the corresponding lunar skin (≃ 5mm). In any soil sample the index of solar wind maturation can be tentatively defined as the degree of saturation of its constituent grains with respect to solar wind effects. But from the previous definition of Δ_{LS}, this index is a complex quantity. Indeed 50 μ-grains are exposed about 50 times longer in the solar wind than the 1 μ-grains. Furthermore ilmenites will reach saturation levels at doses of solar wind ions that are 10 times higher than those observed for the feldspars. Therefore on the lunar surface feldspars of any size mature sooner than ilmenites, and we predict that most of the finest ilmenite grains will not get mature with respect to solar wind effects (ref.1, chapter II.3). In a given soil sample a "feldspar" index of solar wind maturity could then be estimated by measuring the proportion, F, of 1 μ-feldspars showing an amorphous coating of solar wind radiation damaged material. If this proportion is already equal to the saturation value (100 %) one could rely on an "ilmenite" index of maturity determined by measuring the concentration of solar wind implanted species in the finest ilmenites.

IV. THE PAST ACTIVITY OF SOLAR WIND MATURATION

We first propose the following new technique for extracting dust grains in lunar core tubes : 1. several couples of strata showing well defined boundaries have to be identified at various depths in a lunar core tube ; 2. then in the lower stratum of a given couple, grains should be extracted up to a maximum depth, Δ_{LS} ≃ 5mm, from the boundary with the upper adjacent stratum (Δ_{LS} sampling technique). A high proportion of these "Δ_{LS}" grains have been last exposed to the solar wind at a rather well known epoch, which corresponds to the time of deposition of their parent strata on the lunar surface. In addition for the feldspars of any size this "last" solar wind exposure was sufficiently long to both sputter away all previously acquired solar wind effects and to reload at saturation the grain "memory" with fresh effects, which thus reflect the average properties of the solar wind during the last ion implantation episode. The past activity of the solar wind could then be measured as already suggested (2) by using such "Δ_{LS}" feldspars.

V. EFFECTS OF MATURATION ON THE BULK PROPERTIES OF THE LUNAR REGOLITH.

V.1. Albedo of the lunar regolith. The optical properties of the regolith are strongly affected by weathering processes active during the Δ_{LS} and N_S cycling of lunar dust grains. To identify such processes we summarize below our

MATURATION OF LUNAR DUST GRAINS

Bertaut J.L. et al.

albedo studies conducted as follow : the albedo, A, of size fractions and mineral separates extracted from soil samples showing different F index of solar wind maturity (see chapter III) is determined from ∿ 1,300 Å up to 6,500 Å ; then possible correlations between the A values and several characteristics of individual grains from the corresponding sub-samples (measured as described in ref.3) are searched for. Our main results are : The albedo of a mature soil is only similar to that of its finest size fraction ; 2. the finest grains have no "dark" metallic coating (ref. 3, chapter III.1) and they are both enriched in feldspars and depleted in Fe-rich grains such as pyroxenes and glasses (ref. 3, chapter III.5) ; 3. it seems that accretionary particles on the grains are enriched in opaque grains (ref. 3, chapter III.5-B); 4. the finest grains are in fact aggregates of welded dust particles (ref. 3, chapter III.3-A) ; 5. the A values sharply drop when the F index increases above 20 %. On the other hand no clear correlation was found between A and either the proportion of glassy grains in the 5 μ-residues or Al/Si ratios smaller than ≃ 0.5. Therefore the darkening of a lunar soil with an Al/Si ratio $\lesssim$ 0.5 is probably due to a mixture of grain aggregation, solar wind effects, and accretionary particles.

V.2. Concentrations of solar wind implanted species in the regolith. We have applied high resolution mass spectrometric pyrolysis techniques to feldspar and ilmenite targets, implanted with doses of He, C, N, and S ions ranging from 5.10^{15} up to 10^{18} ions/cm^2, at solar wind energies. (A) Helium concentrations versus dose : 1. the retention of He in feldspars, at the temperatures evolved during the ion implantations ($\lesssim$100°C) is quite complex and follows two main courses : first He gets completely lost from the targets and then starts to be retained at a flux value of about 3.10^{16} ions/cm^2 ,which just corresponds to the formation of an amorphous coating of radiation damaged material on the feldspar (this feature probably reflects the transition from a "damage" to a "trapping" diffusion mechanism). But only ≃6% of the incident ions get thus retained up to an amount of ≃3.10^{16} ions/cm^2 ; 2.in ilmenite the He concentration increases up to a saturation value ,Ψ_s(He) ≃3.10^{17} ions/cm^2 ,where about 30% of the incident ions are effectively trapped into the target ; 3.from these features we first predict Ψ_s(Ar) values in the 10084 ilmenite separates analyzed by Eberhardt et al(4) which well agree with the experimental determinations. In addition we reconfirm independently for ilmenite the solar wind sputtering rate (≃0.05 Å/year) that we estimated last year from an electron microscope observation of 1μ-grains. (B) Pyrolysis release patterns of He, C, N, and S :1. all species start to be released at low temperatures (≃150°C) from the targets, with carbon being the most stable species; 2. the patterns exhibit release peaks that are only well defined for carbon (bimodal CO release and trimodal CO_2 release in feldspar) and to a lesser extent in nitrogen (trimodal release in ilmenite); 3.the release peaks broaden and get assymetrical for high doses, and they are generally shifted to higher temperatures when the ion energy increases ;3. carbon is the best"chemical" probe to investigate the properties of the ancient solar wind.

REFERENCES. (1) G.Comstock et al. (1975), this volume .(2) J.Borg et al.(1974), in Solar Wind III,p.68,UCLA. (3) J.Klossa et al. (1975), this volume .(4) Eberhardt P. et al. (1972),Proc.Lunar Sci.Conf. 3rd, p.1821.

D4

KEROGEN STRUCTURES IN RECENTLY-DEPOSITED ALGAL MATS AT LAGUNA MORMONA, BAJA CALIFORNIA: A MODEL SYSTEM FOR THE DETERMINATION OF KEROGEN STRUCTURES IN ANCIENT SEDIMENTS

R.P. Philp and M. Calvin, Department of Chemistry, University of California, Berkeley, California 94720

Presented at the 2nd Environmental Biogeochemistry Conference, Burlington, Canada, April 8 - 11, 1975

ABSTRACT

Laguna Mormona, Baja California, is dominated by an evaporite flat and hypersaline marsh environment which is characterized by algal mats and zones of *Salicornia* and *Distichlis*. Large quantities of organic material are accumulating because of the high salinity, which virtually eliminates the grazing and burrowing activities of metazons which ingest and digest algal and bacterial components of the mat, and also the persistent anaerobic conditions with increasing depth.

Kerogenous material isolated from these algal mats was subjected to chemical degradation treatment. Structural determinations of kerogens from ancient shales have often been thwarted by the complex mixtures of reaction products obtained. It was anticipated that in view of their age, kerogens in these recently deposited algal mats would be structurally less complex. Indeed chemical degradation of this material yielded reaction products less complex than those obtained from ancient kerogens.

The major oxidation products isolated included dicarboxylic acids (C_{11} - C_{22}), α-methyl-branched monocarboxylic acids (C_{13} - C_{20}), normal monocarboxylic acids (C_{12} - C_{30}) and a series of isoprenoid acids dominated by phytanic acid. These results suggest that kerogenous material is present in these algal deposits in a very immature state and it resembles ancient kerogens to a certain extent.

INTRODUCTION

Kerogen is a ubiquitous source of organic carbon on the Earth's surface, containing over three orders of magnitude more organic carbon than all the other organic carbon reservoirs combined (Sackett et al., 1974). Kerogen has been defined in many different ways but in this paper the term kerogen refers to organic material in sediments, shales and rocks which is insoluble in organic solvents. The deposition of plants and animal remains in marine and non-marine environments and the alteration of this debris during subsequent geologic periods by numerous reactions produces a wide variety of carbonaceous materials. Oil shales, kukersite, coorongite and torbanite represent different types of carbonaceous material containing various amounts of kerogenous organic material, from which high yields of oil can be produced at elevated temperatures.

The geochemical, environmental and economical significance of kerogen has led to many attempts to reveal its chemical and physical nature, its method of formation and its ancient precursors. These investigations have been hampered not only by the insolubility and macromolecular nature of the substance but also by the fact that kerogen is not a uniform molecule but most likely a conglomerate of various sub-units of similar, but varying molecular structures. Kerogen is both chemical and biologically very inert and therefore can be used as a "chemical fossil" in sediments. This term was used by Eglinton (1973) to describe various classes of organic compounds which are relatively stable under varying environmental conditions throughout long periods of geological time. An example of kerogen's potential as a chemical fossil was given by Sackett et al. (1974) who analysed amounts and isotopic composition of the organic carbon, which is 95% kerogen, in sediments of the Ross Sea, Antarctica. From this study it was shown that much of the sedimentary organic carbon in the sediments is derived from the rocks being eroded by glaciers on the Antarctic continent.

It is now generally agreed that many oil shales were formed by growth and deposition of algae, and that conditions of sedimentation, compaction and lithification were not too dissimilar from those existing at present. Thorne and co-workers (1964) at the U.S. Bureau of Mines aptly summarized the situation in the following way: "Oil shale was formed by the deposition and lithification of finely divided mineral matter and organic debris in the bottom of shallow lakes and seas. The organic debris resulted from the mechanical and chemical degradation of small aquatic algal organisms." If indeed this is correct then as Cane (1969) has noted, with some surprise, it is unusual that we fail to observe the initial stages of kerogen formation in the contemporary environment. Although as he adds this may be due to the infrequency of the peculiar environmental conditions for favorable vegetal growth. Bradley (1966, 1970) has suggested the reason for failing to observe these initial stages of kerogen formation is due to the fact that the rate of deposition of kerogenous algal ooze is less than 1 mm per century for the dry material. However under certain specialized conditions profuse algal growth does occur and leads to the formation of Coorongite and Balkaschite which represent an early stage in the formation of algal kerogen. Cane (1969) has suggested therefore that Coorongite can be regarded as the "peat" stage in the "coalification" of algal shales and proposed that a study of the nature of this intermediate should provide important information on the nature of the more mature and inert kerogen. Thus from an examination of living algae and recent algal deposits it should be possible to get additional information on the initial stages of kerogen formation.

Consequently a combined study of the soluble and insoluble organic matter present in recently-deposited algal mats from Laguna Mormona, Baja California, has commenced. The soluble lipid material in these mats will be discussed elsewhere in this volume by Eglinton and co-workers whereas we will concentrate on the insoluble or kerogenous fraction of the mats.

Blue-green algae are especially interesting because of their ability to survive in environments which favour the preservation of organic matter such as hypersaline and reducing environments. Contemporary algae are also of interest since the search for the origins of unicellular life on Earth has now reached the point where some of the microstructures referred to as exhibiting "alga-like" and "filamentous" morphologies occur in cherts from South Africa that are older than 3 eons (3×10^9 years, Schopf and Barghoorn, 1967). In order to get a clearer picture as to the origin and nature of these possible organisms it is essential that the nature of modern prokaryotic algae is completely understood in the first instance.

The area at Laguna Mormona is characterised by extreme local variations in sedimentological, geochemical, and biological properties (Vondeer Haar, 1973). Semi-arid climatic condtions and restricted water movement from the ocean have combined to produce an evaporite flat and hypersaline marsh environment. A barrier dune ridge separates the sea from a 100 m wide marsh which is characterised by algal mats. Mucilaginous algal laminae are disrupted by plant roots, early diagenetic growth of aragonite granules and also by dessication cracks. Large quantities of organic matter are slowly accumulating because of the high salinity which inhibits organisms that normally consume and degrade the organic material. ("Algal mat," in general, refers to a cohesive fabric of filaments produced by a community of cyanophytes involving several species combined with different amounts of sediment (Logan *et al.*, 1964). The main mat-producing genera are *Aphanocapsa*, *Aphanothece* and *Entophysalis* among the coccoid, and *Oscillatoria*, *Lyngbya*, *Microcoleus* and *Schizothrix* among the filamentous cyanophytes, although green and red algae, animals and abundant bacteria are often included in the community (Golubic, 1973).)

The intertidal mats examined from Laguna Mormona are characterised by *Microcoleus chthonoplastes* and *Lyngbya aestuarrii*. Sediment binding by *M. Chthonoplastes* is common to a wide variety of environments and no particular mat type can be consistently attributed to it. *Lyngbya aestuarrii* has also been observed in arid subtropic bays such as Shark Bay and the Persian Gulf. In those environments the flat mats covering the floors of tidal pools and tidal channels are laminated structures with different proportions of trapped sediment. The mats often crack during draining, and shrinking of the algal fabric creates polygons of different sizes (Golubic, 1973).

Within the intertidal areas at Laguna Mormona several factors appear to influence the distribution and growth of the mats. They are characterised by the alteration of the water level and exposure to air. A chemical gradient between the limit of the tidal flooding and the evaporite sequence and freshwater influx leads to changes in mat distributions. Differences in mat morphology can be partially accounted for by the amount of sediment deposited over the mats by water movement.

The diverse nature of the organosedimentary structures within the mats can readily be appreciated by summarizing the three general areas of importance that should be considered when examining these structures: (1) the nature and species composition of the algae community; (2) the interaction of the community with major environmental variables, such as location in respect to tides, currents, sedimentation rates, light, drainage and oxygen supply; and (3) the biological dynamics within the community balances between the rates of primary production and bacterial decomposition (Golubic, 1973).

As mentioned above there have been several studies directed at determining the structure of kerogen in ancient shales and sediments, in particular the Green River oil shale (Robinson _et al_., 1953, 1956, 1961, 1963; Burlingame _et al_., 1968a, 1969b; Djuricic _et al_., 1971). Djuricic and co-workers (1971) have also examined a collection of kerogens from ancient sediments of different geographic origin in an attempt to correlate variations in kerogen structures with the history and environment of the shale and thus give valuable information concerning the chemical formation of kerogen. All of these studies have involved degradation of the kerogen, either by oxidation, reduction, hydrolysis or pyrolysis techniques, followed by structural identification of the individual components of the degradation mixtures. From this data it is then possible to attempt the reconstruction of a structure for the original kerogen. The various reagents that have been used in the oxidation studies include chromic acid (Burlingame _et al._, 1969a), hydrogen peroxide (Downs and Himus, 1941), ozone (Jones, 1922) and potassium permanganate (Robinson _et al._, 1953). In the majority of these studies the mixtures of the degradation products obtained have been complex and although it has normally been possible to analyse them by techniques such as computerized-gas chromatography-mass spectrometry after initial fractionation and derivatisation, it is still difficult to construct anything but a preliminary model for the kerogen.

Thus it was anticipated that similar degradation studies of the less mature kerogenous material in these recently deposited algal mats would provide answers to the following questions:

(a) Is there any kerogen-like material in these recently deposited algal mats?

(b) Will degradation studies of young kerogens give information as to the structure and nature of ancient kerogens?

(c) What are the initial steps in the formation of kerogen?

(d) Are changes in complexity of the kerogen observed with changes in depth within these algal mats?

(e) Can kerogen-like material in recently deposited algal mats be used as a chemotaxonomic indicator?

EXPERIMENTAL

(i) Sample collection

A box coring device was used to collect a sample of the algal mats to a depth of approximately eight inches. This was divided into three sections which were transferred to clean glass bottles, capped and subsequently stored at 0°C until required for examination and extraction. No isopropanol was added to the samples immediately after collection as it was felt that changes due to microbial alteration would be minimal if the samples were stored at 0°C.

(ii) Sample preparation

To determine whether or not there was any kerogen-like material produced by degradation and diagenesis of the algal mat and also get an idea of its structure, it was decided in this preliminary study to use an oxidative degradation technique similar to that used by Burlingame and Simoneit (1968b) in their oxidation studies of Green River oil shale kerogens. Almost every published work on the study of kerogen uses a different method for the preparation of the samples making it extremely difficult to compare results of various workers. Forsman and Hunt (1958) proposed an acid digestion of shales with consecutive treatments with HCl, HNO_3 and HF to remove carbonates, pyrites and silicates respectively. Burlingame and Simoneit (1968a) however used zinc dust and 6M HCl to remove sulphides and free sulphur. Lawlor and co-workers (1963) advocated the use of lithium aluminum hydride for quantitative pyrite removal although this results in specific alteration of kerogen functional groups. One important factor to be taken into consideration with all these methods is the possibility of alteration of the organic kerogen by the action of strong mineral acids.

The mineral content of the algal mats was found to be predominantly carbonate and as a consequence of this was treated only with 6M HCl. The method of isolation of the kerogen-like material is shown in Figure 1. After removal of supernatant water the sample was refluxed for 6 hours with 6M HCl. It was centrifuged and the residue exhaustively extracted for a total period of two weeks with the following solvents: toluene/methanol (1:1), methanol and finally toluene.

The supernatant aqueous fraction was set aside for examination of its amino acid, lactone and more polar dicarboxylic acid content at a later date. A more detailed study on the lipids in the total organic extract of the algal mat has been carried out by the Organic Geochemistry Unit at Bristol and these results will be reviewed by Professor Eglinton elsewhere in this volume.

The resulting residue, hereafter referred to as kerogen, was washed with triple distilled water and allowed to dry under vacuo at room temperature. It should be mentioned at this juncture that this residue was not subjected to any further treatment, such as HF treatment or saponification. However subsequent work with this algal mat sample has shown that HF treatment does not have any noticeable effect on the qualitative or quantitative nature of the results described herein (R.P. Philp, unpublished results).

(iii) <u>Oxidative reactions</u>

The kerogen residue was subjected to successive chromic acid oxidations (successive reaction times used: 3 and 6 hours). 4.3 g of kerogen residue (after hydrolysis and extraction) was refluxed 3 hours with 3M chromic acid in 3M sulfuric acid. The residue was then filtered off, washed with water and extracted three times each, first with heptane and then diethyl ether, using ultrasonication to insure thorough extraction (Figure 2). The spent chromic acid solution was also extracted with heptane and then ether. The respective extracts were combined and the acids extracted from them with 1M KOH solution. The acids were recovered and esterified with BF_3/MeOH and the esters of normal acids were separated from those of branched-cyclic acids by clathration with urea. The total acid extracts from the 6 hour oxidation were not subjected to urea clathration.

All analytical gas-liquid chromatography (GLC) was carried out on a Perkin Elmer Model 900 gas chromatograph. The total, normal and branched/cyclic fractions were gas chromatographed on a 10 ft. x 1/16 in. i.d. stainless steel column, packed with 3% Dexsil 300 on Gaschrom Q and programmed from 70 to 280°C at 4°/minute with a helium flow rate of 15 ml/minute. (These same GLC conditions were used throughout this study.) The various components of the mixtures analysed by GLC were identified by their retention times, coinjection of standards, and low resolution mass spectra obtained by combined gas chromatography-mass spectrometry (GC-MS). Combined GC-MS analyses were carried out on a DuPont 492-1 instrument interfaced with a Varian Aerograph Model No. 204 equipped with linear temperature programmer. The column used for the GC-MS analyses was a 30 ft. x 0.03 in. i.d. glass capillary column packed with 1% OV-1 coated on 80-100 mesh Gas Chrom Q. The mass spectral data was acquired and processed using a DuPont 21-094 data system.

RESULTS

The elemental analyses of the algal mat, the kerogen residue and the residues after the various oxidations of the kerogen residue are given in Table 1. The yields of total and acid extracts for all extraction steps are listed in Table 2.

(i) Exhaustively extracted fatty acid fractions

As mentioned a detailed description of the soluble fatty acid extract can be found elsewhere in this volume (Cardosa et al., 1975). However a preliminary examination of this fraction by computerised-gas chromatography-mass spectrometry (C-GC-MS) revealed the normal fraction to predominantly consists of a homologous series of normal saturated fatty acids in the range C_{14} - C_{30} with maxima at C_{16} and C_{26}. The presence of mono-unsaturated C_{16} and C_{18} fatty acids was also detected. Minor quantities of iso- and anteiso-C_{15} acids were observed in the branched/cyclic fractions along with relatively large quantities of a higher molecular weight acid tentatively identified as a C_{32}-triterpanoic acid. In view of the discussion below it is important to note that no dicarboxylic acids could be detected in this fraction by mass fragmentography (Hites and Biemann, 1970) or interpretation of individual spectra.

(ii) 3 hour oxidation

(a) Branched-cyclic fractions

The heptane fraction consisted mainly of a series of isoprenoid mono-carboxylic acids dominated by C_{14}, C_{15}, C_{16}, C_{17}, C_{19}, C_{20} acids (peaks 1, 2, 3, 5, 7, 8 respectively in Figure 3) and their relative distribution is clearly illustrated in the lower gas chromatogram shown in Figure 3. Two other branched aliphatic carboxylic acids (C_{17} and C_{18}, peaks 4 and 6 respectively in Figure 3) were also present. The diethyl ether extract was dominated by a series of mono-methyl branched dicarboxylic acids in the range C_4 - C_8. No aromatic, cyclic or keto acids could be detected in either the heptane or the ether fraction (Figure 4).

(b) Normal fractions

The normal heptane extract was dominated by a series of normal monocarboxylic acids in the range C_{12} - C_{27} with a maximum at C_{16}. A series of saturated unbranched dicarboxylic acids were present in minor quantities in the range C_{13} - C_{27} with a maximum at C_{21}. A second series of saturated branched dicarboxylic acids (range C_{11} - C_{15}, max C_{13}) was also present in minor quantities. The other homologous series identified in this fraction was a series of α-methyl-branched monocarboxylic acids in the range C_{13} - C_{20} with a maximum at C_{17} (Figure 3).

The ether extract was dominated by the more polar unbranched dicarboxylic acids in the range C_8 to C_{27} with a maximum at C_{10}. The acids in the range C_{13} - C_{27} were present only in minor quantities. There was also another major series of saturated branched monocarboxylic acids in this fraction in the range C_8 - C_{13} with a maximum at C_{10}.

(iii) 6 hour oxidation

These fractions were not subjected to urea adduction due to the small quantities of material available. The heptane extract did not give any resolvable peaks by GC and was therefore not further analysed by GC-MS. The ether extract basically consisted of two homologous series. The first was a series of normal acids in the range C_{17} - C_{29} with a maximum at C_{24}, and the other a series of α,ω-dicarboxylic acids in the range C_{12} - C_{24} with a maximum at C_{20}.

DISCUSSION

Before commencing the discussion on the qualitative aspects of these results it is valuable to examine the quantitative results shown in Tables 1 and 2. Firstly from Table 1 it can be seen that after only the 3 hour oxidation the organic carbon content of the kerogen residue has been reduced to 0.86% and is further reduced to 0.54% after 6 hours. Firstly this suggests that there is kerogenous material in this very young algal debris which can be degraded by oxidation. Secondly this kerogen appears less resistant to oxidative degradation than the kerogen from Tasmanian tasmanite. In this case 90 hours were required for the organic carbon content to be reduced to the value obtained for the algal kerogen after a 3 hour oxidation period (Simoneit and Burlingame, 1973).

From Table 2 several important factors emerge, primarily the fact that 58.7 g of material has been removed in the initial hydrolysis experiment. This suggests the presence of large amounts of carbonate in the algal debris. A certain amount of lipid material will also be water soluble and hence not accountable for in this particular work up scheme. The majority of degradation products are produced after the 3 hour oxidation, once again suggesting that the kerogen nucleus is less resistant to oxidation than the kerogens from older shales. The residue remaining after the 6 hour oxidation is predominantly inorganic silicates as shown by X-ray fluorescence measurements.

Again the discrepancy in the materials balance is due to the production of carbon dioxide and water soluble products during the oxidation reactions. It is intended that future experiments will be directed towards permanganate oxidation of the algal kerogens using a method similar to that described by Djuricic and co-workers (1971) which includes removal of the products from the oxidative environment as soon as they are formed.

From these results it appears that the nucleus of this kerogenous material is predominantly a system of cross-linked polymethylene chains of varying chain lengths and various degrees of branching. The kerogen appears to be highly aliphatic, only trace amounts of aromatic acids were detected. It would also appear that condensed to this system on the periphery of the nucleus are unbranched hydrocarbon chains and also isoprenoid chains which when oxidized give rise to the normal and isoprenoid acids respectively. The fact that less time is required to completely oxidize this immature kerogen-like material would suggest that it is not a very highly condensed system as some of the more mature kerogens appear to be. It also suggests that the cross linkages are more susceptible to oxidation than in the older kerogens. In the soluble lipid fractions there are relatively large amounts of cyclic acids, such as the C_{32} triterpanoic acid mentioned above. However only trace amounts of cyclic acids could be detected in the products from the kerogen oxidation. This suggests that the first stage in the formation of the kerogen is the condensation of the functionalized aliphatic lipids to give the cross-linked polymethylene nucleus. The branched dicarboxylic acids and isoprenoid acids could either be condensed on the system at its periphery or alternatively at this very early stage in the kerogen formation could be trapped within the polymethylene matrix prior to condensation. The most probable source of the isoprenoid acids is the phytol side of chlorophyll. Simoneit and co-workers (personal communication) have shown by stereochemical determinations that this indeed is the case with the Green River Shale. The mono-methyl branched dicarboxylic acids from the 3 hour oxidation could possibly arise from the mixture of 7- and 8-methyl-heptadecanes which are known to be the major hydrocarbon constituent of blue-green algae (Han et al., 1968). Again with this young immature kerogen these constituents could be entrapped within the polymethylene matrix and on oxidation give rise to the branched dicarboxylic acids detected in certain fractions.

The results of this preliminary study have indicated that there is kerogenous-like material present in these recently deposited algal mats whose structure bears certain similarities to that of more mature and inert kerogens. However the problem still arises as to the method of formation of the kerogen nucleus in these algal mats. It also remains to be shown whether or not this material is formed by simple condensations of soluble unsaturated lipids known to be present in the algal mats or alternatively as suggested by Oehler and co-workers (1974), intracellular chlorophyll molecules might become grafted onto cellular macromolecules through ester and related linkages and these grafted pigment complexes may become incorporated into the insoluble kerogen fraction.

CONCLUSIONS

(1) On the basis of the degradation products obtained it

can be said that a kerogen like material is formed early on in the deposition of the algal mats.

(2) The main kerogen nucleus seems to consist of polymethylene chains with some cross linking, a small amount of aromatic material, plus unbranched material on the periphery.

(3) The kerogen nucleus is probably derived either from condensation of unsaturated oxygenated compounds or even the chlorophyll moeity.

(4) The presence of relatively abundant quantities of phytanic acid suggests that this originates from the chlorophyll and is ester-linked to the kerogen nucleus.

(5) To expand the idea that kerogen is being formed rapidly a depth study of kerogens from these algal mats will be undertaken to see if the degradation products show increasing complexity with depth.

This initial study has provided some answers to the questions asked above but until further samples are examined it will not be possible to say whether the kerogen can be used as a chemotaxonomic indicator or whether there are changes in its complexity with changes in depth.

Two very important questions which still remain unanswered relate to (1) the method of kerogen formation and (2) the types of linkage that hold the kerogen matrix together. To provide an answer to the first question is difficult. One way to study the short term fate of biolipids is to use ^{14}C-labelled lipids and incubate them in the sediment for a period of several months or more. However in order to observe the incorporation of these biolipids into the kerogen matrix would take an impractical period of time. On the second question one can speculate and say that there are ether linkages or other types of heteroatom linkages in the kerogen which are attacked on oxidation with chromic acid, but one important point to raise here is what are the differences between the linkages in this algal kerogen which appear to be readily attacked by chromic acid and those in ancient kerogens which are fairly resistant to oxidative degradation? Is it merely the type of cross linkage, or is it due to that fact that the older kerogens are more highly condensed and thus potential sites of oxidation more sterically hindered?

It is anticipated that by looking at other algal kerogens and also other slightly older kerogens the answers to these questions and others may become apparent.

ACKNOWLEDGEMENTS

We thank Mr. Stephen Brown for assistance in obtaining the GC-MS data and Ms. Adrienne Ross for typing the manuscript. We also wish to thank Dr. S.W. Awramik and Mr. B. Simoneit for providing critical comments on the manuscript. This work was funded by NASA grant number NGL 05-003-003 and the DuPont data system was purchased on National Cancer Institute contract #NCI-FS-(71)-58.

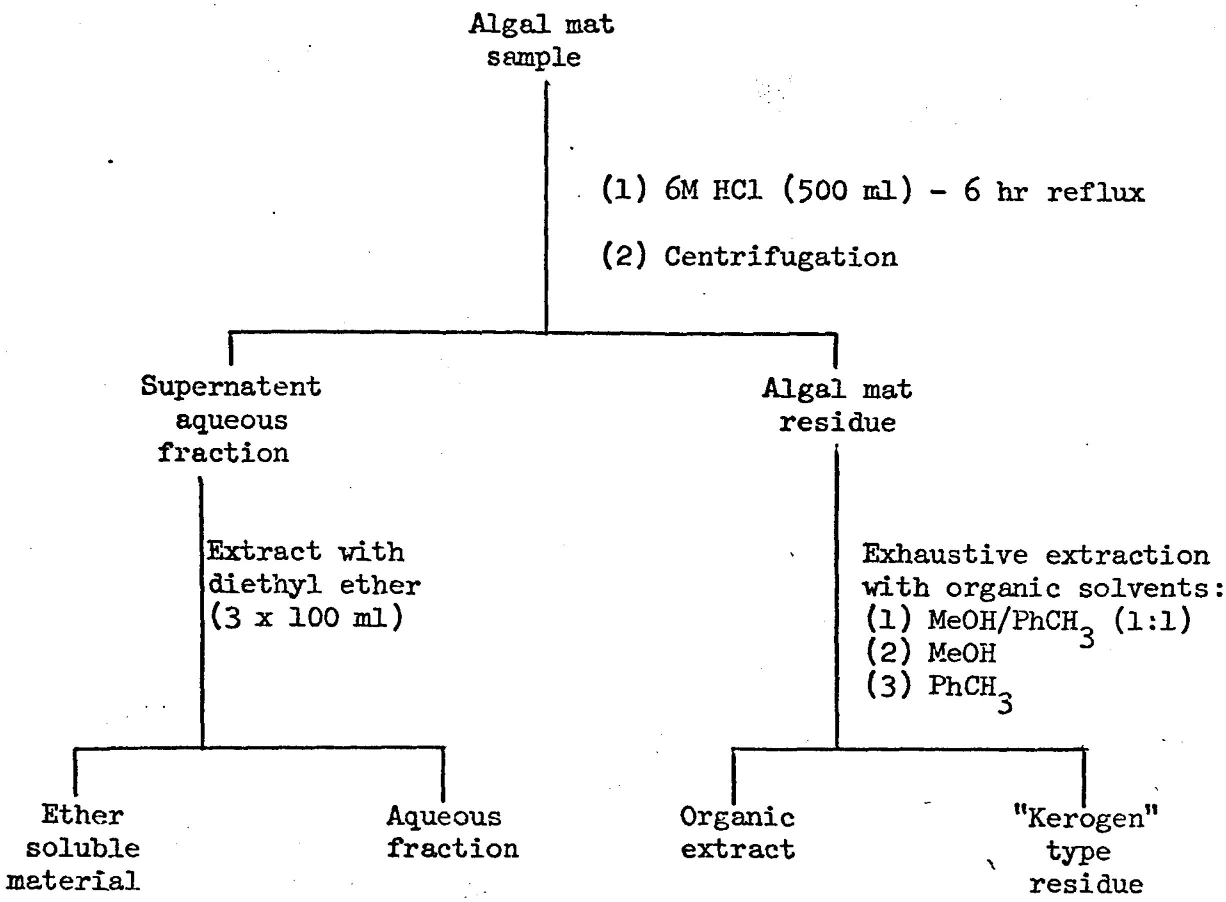
FIGURE CAPTIONS

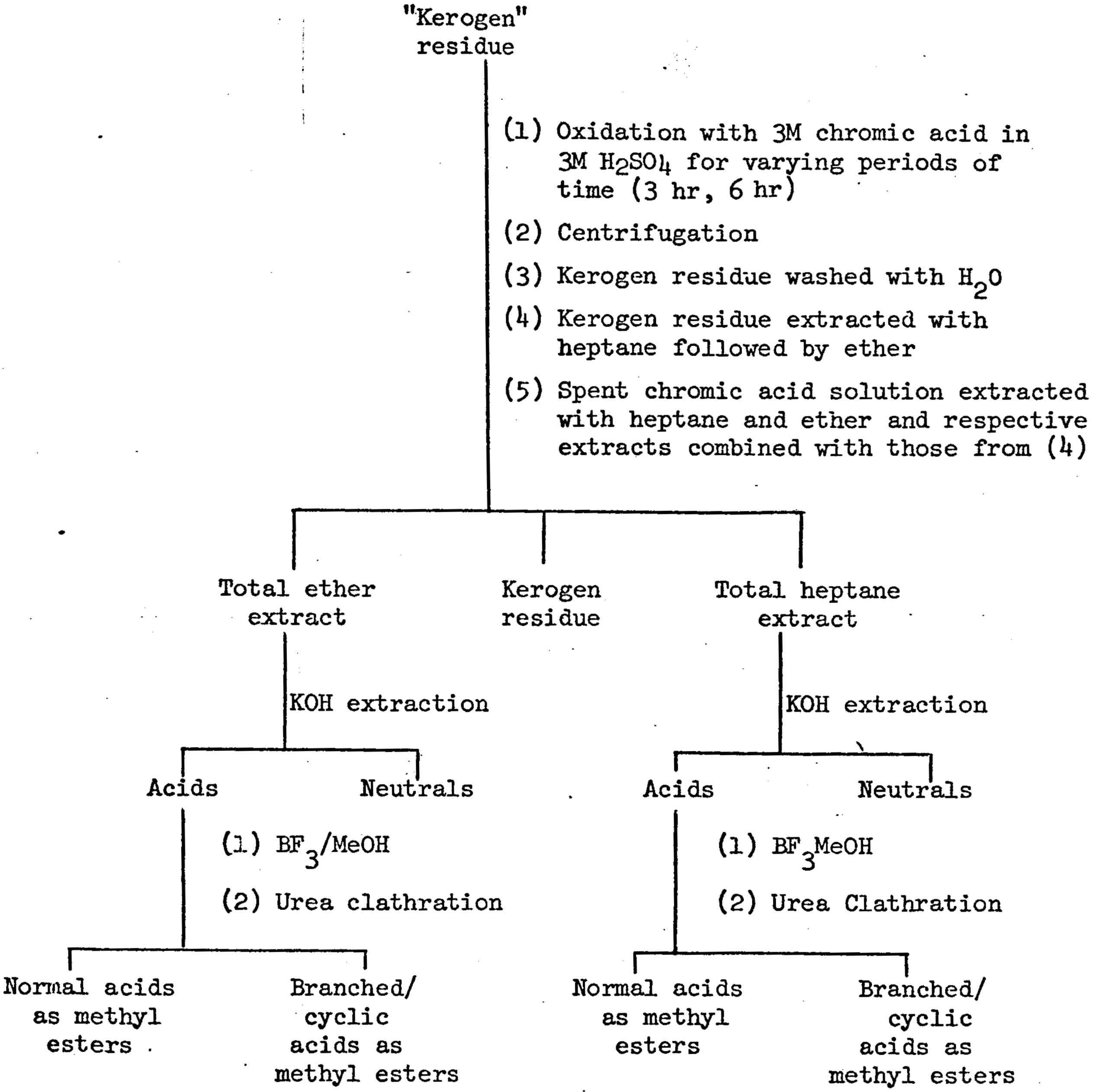
Figure 1. Initial extract and fractionation scheme used to prepare kerogen residues from the recently-deposited algal mat sample.

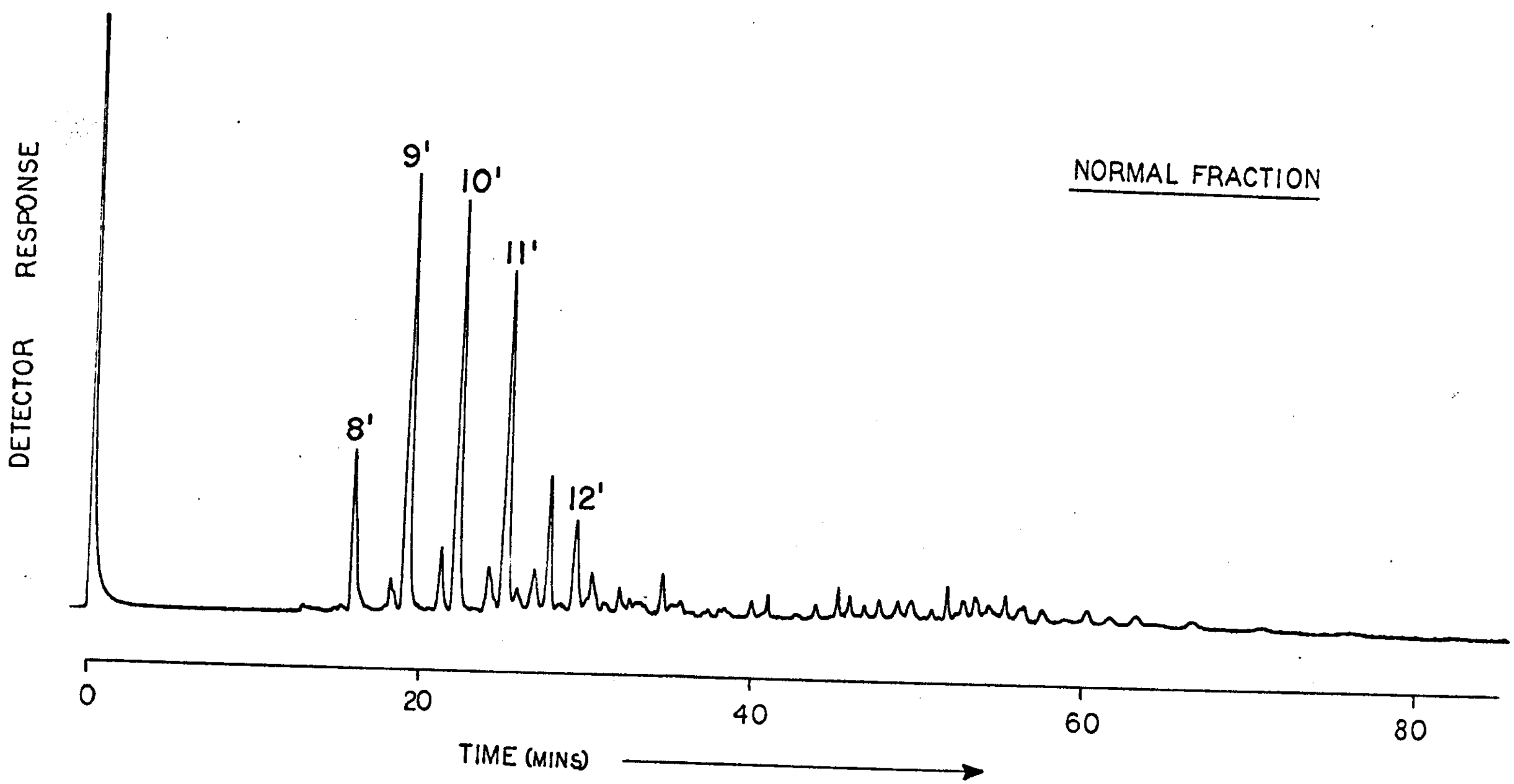
Figure 2. Oxidation scheme used on insoluble kerogen fraction from the algal mats.

Figure 3. Gas chromatograms of the heptane soluble normal and branched/cyclic acid esters isolated from the 3 hour oxidation of the algal kerogen (GC conditions are given in the text). Only the major components in each fraction are labelled in this Figure. In the upper trace the carbon numbers of the normal fatty acids are indicated by the arabic numerals and the monomethyl branched acids indicated by the primed arabic numerals. In the lower trace peak numbers 1, 2, 3, 5, 7, 8 correspond to isoprenoid acids and numbers 4 and 6 to branched acids whose structures have not been assigned as yet.

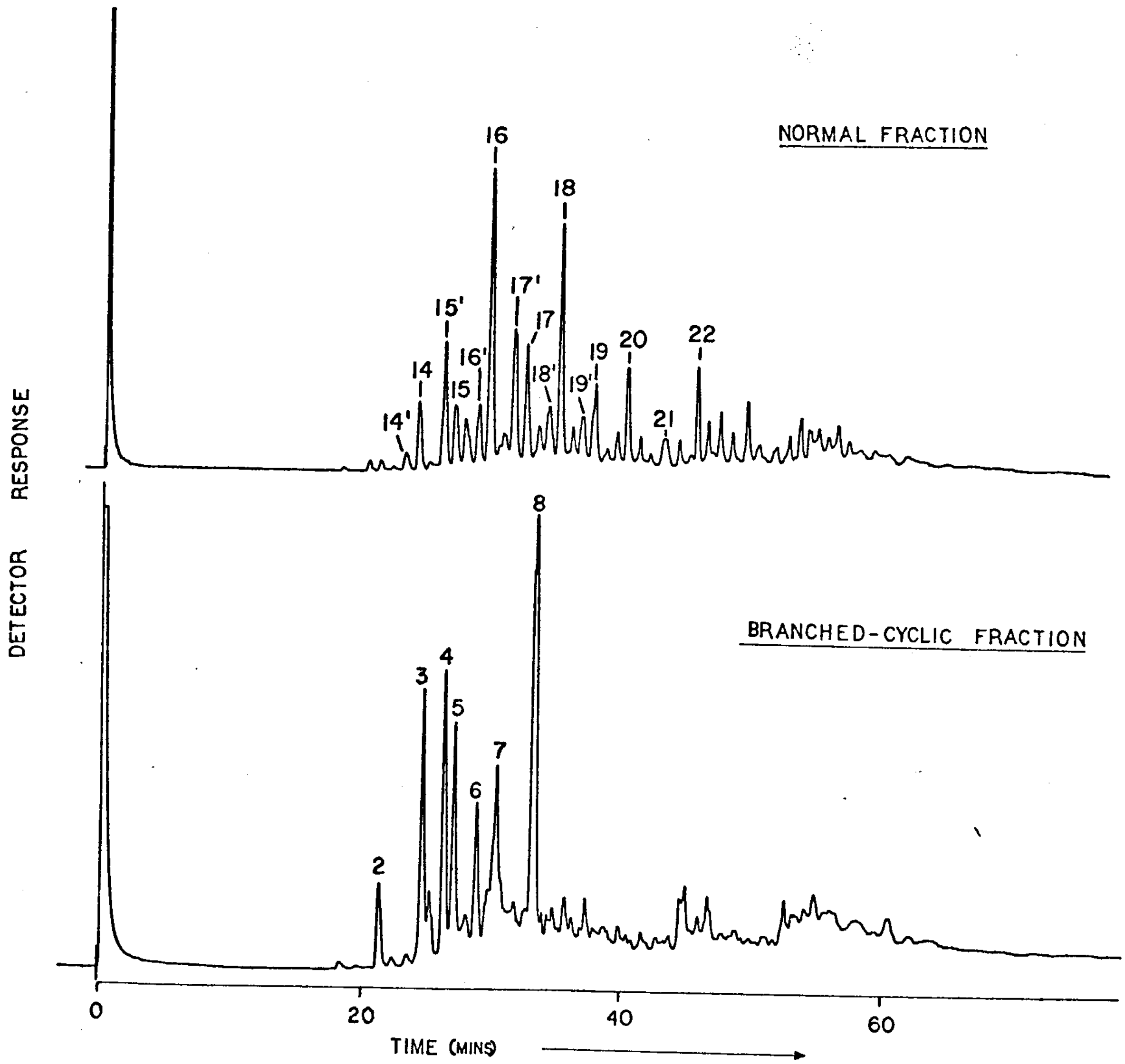
Figure 4. Gas chromatograms of the ether soluble normal acid esters isolated from the 3 hour oxidation of the algal kerogen (GC conditions are given in the text). The primed arabic numerals refer to the carbon number of the normal dicarboxylic acid esters.


Algal mat sample
(1) 6M HCl (500 ml) - 6 hr reflux
(2) Centrifugation
Supernatent aqueous fraction
Algal mat residue
Extract with diethyl ether (3 x 100 ml)
Exhaustive extraction with organic solvents:
(1) MeOH/PhCH$_3$ (1:1)
(2) MeOH
(3) PhCH$_3$
Ether soluble material
Aqueous fraction
Organic extract
"Kerogen" type residue


"Kerogen" residue
(1) Oxidation with 3M chromic acid in 3M H_2SO_4 for varying periods of time (3 hr, 6 hr)
(2) Centrifugation
(3) Kerogen residue washed with H_2O
(4) Kerogen residue extracted with heptane followed by ether
(5) Spent chromic acid solution extracted with heptane and ether and respective extracts combined with those from (4)
Total ether extract
Kerogen residue
Total heptane extract
KOH extraction
KOH extraction
Acids
Neutrals
Acids
Neutrals
(1) BF_3/MeOH
(2) Urea clathration
(1) BF_3MeOH
(2) Urea Clathration
Normal acids as methyl esters
Branched/ cyclic acids as methyl esters
Normal acids as methyl esters
Branched/ cyclic acids as methyl esters



XBL 7411-8574



XBL 7411-8575

Table 1. Elemental analyses of algal mat and kerogen residues at the various degradation steps.

	Dried algal mats (%)	Kerogen residue (%)	After 3 hour oxidation (%)	After 6 hour oxidation (%)
C	9.3†	10.31*	0.86*	0.54*
H	1.8	1.59	1.11	0.82
N	0.38	0.74	0.10	0.03
S	0.12	0.52	0.45	0.08
Residue	67.9	79.8	88.3	93.5
O**	20.5	7.04	9.18	5.03

† Total carbon content

* Total organic carbon content

** By difference

Table 2. Extract yields from extraction and oxidations of the algal mat sample.

Sample	Totals (extracts or residues)	Acid fractions (mgs) Heptane soluble	Ether soluble
1. Dried algal mat	63 g	—	—
2. Hydrolysed and exhaustively extracted mat (i.e, kerogen residue)	4.3 g	—	—
3. Soluble lipid extract from hydrolysed (1)	54 mg	39*	
4. 3 hour Oxidation of (2)	149 mg	18	95
5. 6 hour Oxidation of residue from (4)	24 mg	2	9
6. Residue after 6 hour Oxidation	1.9 g		
7. Total solvent solubles and residues recovered	1.127 g	20	104

* Total solvent soluble acids

REFERENCES

Bradley, W.H., 1966, Tropical lakes, copropel and oil shale. Bull. Geol.Soc.Amer. 77, 1333-1337.

Bradley, W.H., 1970, Green River oil shale - concept of origin extended. Bull.Geol.Soc.Amer. 81, 985-1000.

Burlingame, A.L., Haug, P.A., Schnoes, H.K. and Simoneit, B.R., 1969a, Fatty acids derived from the Green River formation oil shale by extractions and oxidations - a review. In: Schenck, P.A. and Havenaar, I. (eds.), Advances in Organic Geochemistry 1968, pp. 85-129, Pergamon Press, Oxford.

Burlingame, A.L. and Simoneit, B.R., 1968a, Isoprenoid fatty acids isolated from the kerogen matrix of the Green River formation (Eocene). Science 160, 531-533.

Burlingame, A.L. and Simoneit, B.R., 1968b, Analysis of the mineral entrapped fatty acids isolated from the Green River formation. Nature 218, 252-256.

Burlingame, A.L. and Simoneit, B.R., 1969b, High resolution mass spectrometry of Green River formation kerogen oxidations. Nature 222, 741-747.

Cane, R.F., 1969, Coorongite and the genesis of oil shale. Geochim. Cosmochim.Acta 33, 257-265.

Cardosa, J., Brooks, P., Eglinton, G., Goodfellow, R., Maxwell, J.R. and Philp, R.P., 1975, Lipids in recently deposited algal mats at Mormona, Baja California. This volume.

Down, A.L. and Himus, G.W., 1974, A preliminary study of the constitution of kerogen. J.Inst.Pet. 27, 426-444.

Djuricic, M.V., Vitorovic, D., Andresen, B.D., Hertz, H.S., Murphy, R.C., Preti, G. and Biemann, K., 1971, Acids obtained by oxidation of kerogens of ancient sediments of different geographic origin. In: von Gaertner, H.R. and Wehner, H. (eds.), Advances in Organic Geochemistry 1971, p. 305-321, Pergamon Press, Oxford.

Eglinton, G., 1973, Chemical Fossils: a combined organic geochemical and environmental approach. Pure and App.Chem. 34, 611-632.

Forsman, J.P. and Hunt, J.M., 1958, Insoluble organic matter (kerogen) in sedimentary rocks. Geochim.Cosmochim.Acta 15, 170-182.

Golubic, S., 1973, The relationship between blue-green algae and carbonate deposits. In: Carr, N.G. and Whitton, B.A. (eds.), The Biology of Blue-Green Algae, p. 434-472, University of California Press, Berkeley and Los Angeles.

Han, J., McCarthy, E.D., Van Hoeven, W., Calvin, M. and Bradley, W.H., 1968, Organic geochemical studies II. A preliminary report on the distribution of aliphatic hydrocarbons in algae, in bacteria and in a recent lake sediment. Proc.Nat.Acad.Sci. USA 59, 29-33.

Hites, R.A. and Biemann, K., 1970, Computer evaluation of continuously scanned mass spectra of gas chromatographic affluents. Anal.Chem. 42, 855-860.

Jones, C.L., 1922, The problems of the shale oil industry. Chem.Met. Eng. 26, 546-553.

Lawlor, D.L., Fester, J.I., and Robinson, W.E., 1963, Pyrite removal from oil shale concentrates using lithium aluminium hydride. Fuel 42, 239-244.

Logan, B.W., Rezak, R. and Ginsburg, R.N., 1964, Classification and environmental significance of algal stromatolites. J.Geol. 72, 68-83.

Oehler, J.H., Aizenshtat, Z. and Schopf, J.W., 1974, Thermal alteration of blue-green algae and blue-green algal chlorophyll. Am.Assoc. Petroleum Geologists Bull. 58, 124-132.

Robinson, W.E., Cummins, J.J. and Stanfield, K.E., 1956, Constitution of organic acids prepared from Colorado oil shale based on n-butyl esters. Ind.Eng.Chem. 48, 1134-1138.

Robinson, W.E., Heady, H.H. and Hubbard, A.B., 1953, Alkaline permanganate oxidation of oil-shale kerogen. Ind.Eng.Chem. 45, 788-791.

Robinson, W.E. and Lawlor, D.L., 1961, Constitution of hydrocarbon-like materials derived from kerogen oxidation products. Fuel 40, 375-388.

Robinson, W.E., Lawlor, D.L., Cummins, J.J. and Fester, J.I., 1963, Oxidation of Colorado oil shale. Bureau of Mines, Rept. of Invest. 6166, 33 pp.

Sackett, W.M., Poag, C.W. and Eadie, B.J., 1974, Kerogen recycling in the Ross Sea, Antarctica. Science 185, 1045-1047.

Schopf, J.W. and Barghoorn, E.S., 1967, Alga-like fossils from the early precambrian of South Africa. Science 156, 508-512.

Simoneit, B.R. and Burlingame, A.L., 1973, Carboxylic acids derived from Tasmanian tasmanite by extractions and kerogen oxidations. Geochim.Cosmochim.Acta 37, 595-610.

Thorne, H.M., Standfield, K.E., Dinneen, G.U. and Murphy, W.I., 1964, Oil Shale technology: A review. Information Circular 8216 U.S. Bureau of Mines, Washington.

Vondeer Haar, S.P., 1973, Evaporite environment at Laguna Mormona, Pacific coast of Baja California, Mexico. Geol.Soc.Amer.Abs. with Programs 5, 117-118.

D5

(19)

The Costs of Efficiency

Implications of Educational Technology*

IDA R. HOOS

This paper traces the impacts of technology on education and illustrates the steady progression, always rationalized by the ethos of efficiency, from teacher aids to computer-aided instruction, from data-processes to facilitate record-keeping to omnivorous management information systems. The point is made that at present a management syndrome pervades education planning at all levels. This has created an environment hospitable to the family of techniques related to systems analysis, with its companions and components, cost/benefit measures, program budgeting. Quantitative concepts such as productivity, performance, and accountability have rendered the notion of academic excellence virtually obsolete. After a review of the current state-of-the art, the paper finds that industry-style efficiency may be more costly than beneficial to education.

Education, to an extent greater, perhaps, than any other societal institution, has been vulnerable to invasion, if not takeover, by "technology," in its myriad manifestations. Reason for this susceptibility is readily apparent. Forever entangled in the mesh of causality

*Editor's note: The *Journal* periodically will publish an article which in its broad focus on the institution of education also addresses those with postsecondary interests. This paper is presented as having such an orientation.

IDA R. HOOS *is a research sociologist at the Space Sciences Laboratory, University of California, Berkeley.*

that makes it at one and the same time creator, curator, reflector, and transmitter of the values of the society, education is perforce in constant flux. For this seeming instability it is always open to criticism for not being responsive or relevant, both of which are almost unattainable, especially during periods of social upheaval. Moreover, the high visibility of its processes and products render education a likely target for the axe-wielding of every politician from county to Congress and for axe-grinding by psychologists, sociologists, economists, and many others possessed of a zeal to reform or "innovate." Always a proper outlet for the "community participation" impulse, education is bossed by school boards made up of businessmen, buffeted by frustrated fringe activists from various arenas, and used as a steppingstone by politicians, more thirsty for advancement of their own personal power than for that of public knowledge.

Echoing the tenor of the times, education has abandoned the personnel practices that made teaching a catchall for spinsters and sissies, that arbitrarily created "counselors" out of teachers who could not cope in the classroom, and that perversely rewarded competence by transferring it from classroom to administration. The organization of teachers into strong unions, mobilization of women into egalitarian movements, and a new set of social mores have brought about fundamental changes, too numerous to recount and too obvious to review here. What can be pointed out is that concomitantly goals for and expectations from education have undergone transformation.

In the early days of public education in the United States, the prime objective was inculcation of the virtues of self-control and self-discipline. To preserve and promote the social order were considered proper activities for the school.[1] Industrialization and immigration in the latter part of the nineteenth century shifted emphasis to the "escalator" potential of education. Representing an avenue for upward mobility, education became synonymous with opportunity, a stereotype that has evinced remarkable resiliency in the face of repeated attacks over the past 35 years. Changing expectations have caused shifts to, away from, and back again toward vocational education, all within the span of two generations. The revival of prayer and religion in the schools suggests a return to the good old fundamentalist values of yore. Darwin has been banished from science courses; there is nostalgic yearning for the teaching of morality.

[1]R. J. Wiebe, "The Social Functions of Public Education," *American Quarterly*, 21 (Summer, 1969), 147-64.

With education at once a reflection of and a response to the prevailing societal codes, certain adaptations of the current ethos of efficiency were bound to occur. The proposition being that teachers are overworked, and the corollary, that any device or procedure that can be called a teachers' aid is beneficial, instant rationalization and justification have been on tap for a succession of linked phenomena that began with the introduction of "objective tests," includes the multifarious uses of computers for reporting, teaching, and management, and brings to education the space age concepts and tools popularly called systems analysis, which encompasses cost/benefit analysis, program budgeting, performance evaluation, and even Delphian prediction.

"EFFICIENCY" THROUGH TECHNOLOGY

Long before automation in the more tangible and expensive form of electronic data-processing entered the schools, efficiency manifested itself in the form of self-administered, true-false, or multiple-choice ("cafeteria") tests. Because they were machine-processable, they qualified as teachers' aids and soon became the basis for all mass-grading efforts. The College Entrance Examination Board abandoned essay-type questions in favor of SATs (Scholastic Aptitude Tests), grading of which is totally automated. The student's total capacity to communicate via the written word was couched in the one short biographical essay so fraught with life-or-death importance as to paralyze the most talented. Qualifying examinations for graduate work (Graduate Record Examination Tests) are of the same type as are also those intended for specialized curricula such as medicine and law. These instruments, designed to screen applicants with dispatch, have satisfied the criterion of speed and thus fulfilled the expectation of efficiency but not without costs. We now are beginning to perceive that standardized tests, devised primarily as an aid to convenient classification by teachers, deans, and administrators, have exerted a feedback effect on course content and curriculum planning. Many educators recognize the inadequacy of this type of testing as an index to scholastic ability; students with the greatest intellectual promise being at a particular disadvantage with the limited, mechanical format.[2] As

[2] Ida R. Hoos, *Systems Analysis in Public Policy* (Berkeley, Calif.: University of California Press, 1973), pp. 162 ff.

we begin to assess costs, we realize that the devil-take-the-hindmost calculus that valued the shortcut neglected deep-seated consequences. The bargain was, indeed, Faustian.

With the art of communication reduced to making check-marks, Johnny cannot write. He gets through grade and high school without having learned to articulate his thoughts on paper. If he enters college, "bonehead English" cannot remedy the accumulated deficiency created by 12 years of neglect. Throughout his college career, his inability to express himself in writing is a handicap. Students, even Ph.D. candidates, have been known to freeze into a kind of numbness at the stage of dissertation writing. As any personnel officer or teacher in adult education classes can attest, the Johnnies and Janes who have to fill out as much as a one-sentence item on an application form are struck with a kind of petrification. They simply cannot write and hence resort to the most desperate face-saving. The incidence of lost or broken reading-glasses is impressive; the transparence of the pretext to carry the form home, pathetic. In occupations such as police work, where report-writing is crucial, this lack of ability can cause serious internal operational problems.

Before the advent of the computer, the report card served as a communication link between school and home. Parents were required to sign it, and the "Remarks" column invited dialogue, if only a fairly superficial sharing of observations, about the child. At semester's end, the card represented an interesting cumulative summary. No perfunctory matter, this, for the report card was the high point in the term. The IBM machined printout has changed all this. Like a beer can, the card is now one-way and disposable. Teachers' comments in coded form convey ambiguity, the numeral 5 meaning "assignments not turned in and/or properly prepared," and 8, "misconduct and/or poor attitude." Response from parents is neither invited nor expected; the system simply cannot accommodate them. Paradoxically, while pupil-processing efficiency increases, school administrators hire consultants to help them overcome alienation and apathy. Encounter groups and public relations gimmicks, such as "Open House Night," "Black History Week," and "Ecology Week," are supposed to enliven parental interest, often deadened beyond recall.

Thanks to "flexible classroom scheduling," a computer program developed by Stanford Research Institute and used widely in the United States and Japan, the few options allowed high school seniors for honors performance have been withdrawn. Where previously the students were privileged to select a particular teacher or aspect of English as reward for earlier endeavor, now the system makes any such deviation prohibitively expensive. The same schools that buy such programs in the name of efficiency alienate, and thus effectively dampen enthusiasm, but expend time and resources on workshops solemnly devoted to motivation. Teachers and counselors, under pressure to deliver to the system the information it constantly demands and in the form it requires, complain that they have little time for personal contact with the students. Suggested as a remedy are group counseling sessions, utilizing closed-circuit television, or computer programs set to conduct counseling interviews.[3] In the later years of the lives of these students, when "relating" becomes a yearning necessity, *ersatz* primary groups are formed, with political ideology, religion, sex, or economy variously serving as the basis for association. Sometimes the course taken is one of total rejection—of dropping out. All the while, the analysts seriously ponder *anomie* and alienation in journal articles and at annual conferences in Vienna.

"TECHNOLOGY" AS NOSTRUM

Always the public whipping boy, education has been accused by "experts" of harboring a 30-year lag between "innovation and widespread adoption of the innovation." A Congressional report quotes an industrialist who happens also to be a school board president as saying, "the aircraft industry would go out of business in two years if it changed as slowly as education."[4] Not only is this statement presumptuous nonsense, but the implied ascription to it of authoritativeness downright sophomoric. To begin with, "innovativeness," the "buzz" word since the early sixties, does not stand as an unchallenged *plus.* Although the *status quo* may not represent the epitome of perfection, there is nothing intrinsically good about *change.* Innovation, under certain circumstances, can cause havoc. In the second place, the aircraft industry is an unfortunate selection as model in that (1) it feeds sumptuously at the public trough from R & D to operating subsidy,

---

[3]Harry F. Silberman, "Applications of Computers in Education," paper presented at the American Management Association Conference, New York, August 9, 1967, System Development Corporation SP-2909/000/01, Santa Monica, California, p. 8.

[4]"Automation and Technology in Education," a Report of the Subcommittee on Economic Progress of the Joint Economic Committee. Congress of the United States, 89th Congress, Second Session, August 1966, p. 3.

and (2) its profligate mismanagerial practices are shielded from view by the hanky panky that accompanies high-level transfers of allegiance in the game of musical chairs played by retired "brass," government bigwigs, and corporate potentates. Education does not enjoy the comfort of knowing that its failures will be rewarded by prompt Congressional appropriations. Education, always maligned, must write a success story. And, using crocodile tears of sympathy as leverage, technologists make the most of the need to create a favorable image and such occasions as Congressional hearings to promulgate their wares.

For example, in the Report put out, significantly, by the Joint Economic Committee and not the Education Committee, commiseration pours forth over the orders of magnitude by which "the amount of knowledge to be communicated during the process of education" is increasing. Stupefying statistics are paraded:

> There have been estimates that as much technical knowledge will be developed in the next 30 years as has been accumulated in the entire past history of mankind. In this country alone, we produce approximately 25,000 technical papers every week, along with 400 books and 3,500 articles.[5]

Then follows the predictable polemic: "Facing such prospective requirements, our educational system is threatened by a decline in *relative effectiveness* unless *improved productivity* can be brought to its aid." Emphasis is here added to some of the foregoing phrases to call attention to the classic elements of the sales campaign being mounted. The sly equating of the current information explosion with knowledge, the onus of ingesting this superabundance, subtly assigned to the educational system, "effectiveness" and "improved productivity" slipped in as terms generally accepted and clearly understood—all of these serve as preamble to the "pitch."

The solution put forward is, of course, technological aids to education. And these are listed to include "educational television, both open and closed; video tape; computerized instruction; the use of computers for student testing, guidance, and evaluation, and the storage, retrieval, and distribution of information; programmed courses of instruction, teaching machines, particularly the 'talking typewriter'; the use of microfilm and microfilm viewing equipment; and language laboratories." Many of these items, presented as assets, actually exem-

[5] Ibid., p. 4.

plify egregious costs of efficiency, as will be noted later in this paper as preparation for the discussion with which we will mainly be concerned—that of systems analytic techniques and their effects on education. As we have seen in the remarks cited above, there is a definite progression in the forensic; it moves seductively from technological aids for teachers to techniques that mold the very philosophy of education. The systems approach, represented to the Congressional committee as "the creative combination of a variety of skills and devices to produce desired results, an approach that is proving highly successful in the military sphere," becomes *the* prevailing paradigm, the way education will be conceived, contrived, and executed.[6]

COMPUTER-AIDED INSTRUCTION

Computer-aided instruction has its roots in programmed instruction, a 1950 phenomenon which received considerable impetus from the work of Skinner, the latter-day developer of behaviorism and the exploiter of all its associated principles once, but not for all, in deep disrepute.[7] By 1968, several thousand students ranging from elementary school to university level had received a significant portion of their instruction in at least one subject area under computer control. In projects developed at Stanford University under a grant from the Carnegie Foundation and, later, Title IV funds, elementary school children designated as "culturally disadvantaged," were brought by bus to receive instruction from a program in initial reading and mathematics. It works this way:

> The first thing the student does is peck out his name on the typewriter. The computer then searches its memory to determine what the student did on his last lesson and what his lesson should be this time. If it is a reading drill, for example, the computer may display a word on the screen at the same time a recorded voice pronounces the word in the earphones. After several words are simultaneously displayed and pronounced, they are displayed on the screen and the recorded voice asks the student to pick out one. This can be done either on the typewriter or with the light pen. If the student selects the right answer, the recorded voice in his earphone says, "Yes, that's correct."[8]

[6] Ibid., p. 5.
[7] B. F. Skinner, *The Technology of Teaching* (New York: Appleton-Century-Crofts, 1968).
[8] John Rhea, "1968 Seen Critical for Computer Education," *Aerospace Technology*, (January, 1968), p. 20.

Dollar costs of the Stanford, and all systems, which involve capital investment in hardware, are high. The addition of sophisticated audio and graphic components increases costs substantially, with greatly accelerated utilization the only likely means to reduction of cost-per-student-hour. It is on the assumption of a tremendous growth in computerized education that the designers of the Stanford system, which has been replicated, emulated, and expanded elsewhere, with generous federal and foundation funding, confidently calculate lowering of future per unit costs as compared to that associated with the "approximately 3,000 students . . . processed daily" in the Stanford project.[9] Not at all surprising, the same architects of this pupil processing project deplored the absence of concrete measures of efficiency. In a wishful projection of their assembly-line notion of education, they urged a definition of "some standard unit, some 'erg' of learning and forgetting."[10]

The current preoccupation among experts in educational technology as to the distinctions between and relative merits of "computer-aided" and "computer-managed" instruction serves only to obscure the basic issues, not to eliminate them. Technical in emphasis and advocative in conclusion, the debate ignores the possibility that this substitution of a "learning facilitator" for a real-life teacher[11] and the notion of "behavioral engineering," with attendant concepts, such as "stimulus control" and "contingency management,"[12] violate basic psychological principles of learning, as exemplified in an experiment conducted by Gordon W. Allport some years ago.[13] In it, 250 college students were asked to write down three vivid memories of their eighth grade school work and to indicate the kind and degree of their active participation. Were they reciting, performing, arguing, talking? Or were they listening or watching without overt involvement? Allport found that three-quarters of the responses recalled situations in which the subject himself was actively participating. This experiment, corroborated in real-life experience, has been repeated many times.

As for the "culturally deprived children," who are the prime targets for hard sales campaigns for educational technology, a socially

---

[9]R. C. Atkinson and H. A. Wilson, "Computer-Assisted Instruction," *Science*, (October, 1968), p. 73.

[10]Ibid., p. 76.

[11]Perry E. Rosove, "Toward Education in Real Time," paper prepared for session on "Real Life Information Systems and the Public Interest," presented at Fall Joint Computer Conference, San Francisco, Calif., Dec. 9-11, 1968, SP-3140, Santa Monica, Calif. System Development Corporation, August 7, 1968, p. 20.

[12]Harvey J. Brudner, "Computer-Managed Instruction," *Science*, 162 (1968), 962-76.

[13]Gordon W. Allport, *Personality and Social Encounter* (Boston: Beacon Press, 1960) p. 185.

isolated, dehumanized learning environment is possibly the worst kind of approach to their special needs. It may be that a very significant factor in their deprivation is the predominant role played by an electronic box in their lives. Baby-sitter, beguiler, educator, supersalesman, mentor, television, with its one-way stimulus, has so conditioned them that they consider education all play and no work. Taught only to expect a kind of entertainment and deprived of the stimulation that comes into the learning situation from interaction with teacher and with other students, the "culturally deprived" pupils reach high school not only unable, but not interested in learning how, to read. Instant relevance must be provided in special courses at the college level in order to retain them and thus preserve the facade of "equal opportunity," "affirmative action," or whatever seems to have "retention value" for all students and constitutes open sesame to federal funding. Thus, the University of Maine offers a course in wine-tasting, a Florida college gives credit to "film going"; at the University of California a class, Forms of Folklore, Anthropology 159, collects and analyzes limericks, graffitti, dirty and ethnic jokes, myths, and legends. Students were asked why the class was so popular. One girl, who admitted signing up because the scheduling was convenient and the five credits useful, replied, "It's interesting. We do more laughing than anything else." Another student was glad he got into the class because "I dig telling jokes, and I might be able to pick up some new ones from him (the professor)."[14]

Computerized education has been criticized by psychologists for taking into account only a small segment of the learning environment and thus skewing the process in ways that are ultimately destructive. Criticism has also been raised on social grounds, the contention being that "noneducation experts" are dominating the field. The aggressive invasion by hardware merchants, software dealers, engineers, operations researchers, systems analysts, and the like portend a total and self-perpetuating technological takeover. The possibility has been mentioned that the $50-billion-a-year education market, which Xerox, RCA, General Learning Corporation, Raytheon, Lockheed, and others are vigorously exploiting, could lead to an "education-industrial complex," in which large corporations would determine what should be taught. As yet unmentioned is the obvious link between programmed instruction and propaganda, although apparently the White House Office of Telecommunications Policy has already found "potential

[14]"Jokes, Myths, Fill Folklore Class," *The Daily California*, February 8, 1974.

abuse" in the use of audio-visual materials, in that public service messages could become vehicles for "thought implantation."[15] Skinner and his followers having already reduced education to a low form of conditioning, the likelihood of determined brainwashing is not at all far-fetched.[16]

SYSTEMS TECHNIQUES AS APPLIED TO EDUCATION

To belabor the point that education at all levels from local to federal is mismanaged is tantamount to the reverse of an attack on motherhood. Everyone can and does criticize education, its shortcomings, and failures. And, no doubt, with as much, and probably more, reason than to censure other public institutions. Heir to personnel practices that put basketball coaches into administrative position just because those jobs were reserved almost exclusively for males, caught on the pendulum that swings from the classical to the vocational, from the scholarly to the relevant, buffeted by the forces of social change that cast the schools as prime instrument of ethnic, class, and cultural integration, assailed by ideologues and idealists, pragmatists and politicians, parents and professionals, education provides a slow-moving target vulnerable and undefensible.

That systems analysis and all its attendant quantified techniques touted to improve managerial efficiency should emerge as a promising panacea ubiquitously applied seems only logical. The systems approach has appeal, especially in its philosophic guise of "viewing a problem or situation in its entirety with all its ramifications, with all its interior interactions, with all its exterior connections and with full cognizance of its place in its context."[17] Systems analysis is, however, not a merely philosophic notion but an operational actuality. In fact, it is used interchangeably with *operations analysis* and, as such, refers to a specific technique that consists of constructing a mathematical model of an operation. This model is a set of equations which presumably will allow one to calculate changes in the outputs of an operation when specified inputs have been changed, the purpose being to optimize some aspect of

[15]Les Brown, "Audio-Visual Spending by U.S. cited as 'Potential For Abuse,'" *The New York Times*, February 12, 1974.

[16]B. F. Skinner, *Beyond Freedom and Dignity* (New York: Alfred A. Knopf, 1971).

[17]Alexander M. Mood, "On Some Basic Steps in the Application of Systems Analysis to Instruction," *Socio-Economic Planning Science*, 1 (1967), pp. 19-26.

the outputs. The systems analysis does not attempt to be complete in its mathematical model, but rather "to be judiciously selective by including significant factors and omitting or aggregating minor factors in order to keep the model to manageable proportions for purposes of the optimization."[18] This statement, so guilelessly descriptive, contains the very hidden agenda that undermine the credibility of these techniques as "scientific," "objective," and "rational." "To be judiciously selective" is to impose a judgment, which reflects the *Weltanschauung* of the analyst. The distinction between "significant" and "minor" factors is an expression of preference, bias, and, possibly, ignorance. "To keep the model to manageable proportions" discloses the current disease, once called "modelitis" by Kahn; preoccupation with the model often causes neglect of the real-life situation.

History and happenstance have encouraged the wide and ever-increasing use of the techniques. Ours being the Technological Era, the Space Age, the systems approach has become the "dominant paradigm,"[19] of our time and, as such, is being applied to every facet of our lives—health, welfare, transportation design, land use, education, and even the future of society. Moreover, systems analysis epitomizes the Space Age. When the overworked cliché, "A nation that can put a man on the moon" is evoked as evidence of technical prowess, reference is being made to the systems methods used with such singular success in space exploration.

Another pillar of prestige stems from the World War II experience when, under the name of Operations Research, the techniques were credited with having won the Battle of Britain. Secretary of Defense Robert S. McNamara, with his Rand associates, added a dazzling chapter to the saga of systems analysis. That the wizardry was short-lived, that subsequent disclosures by Congressional committees found the systems approach to be directly responsible for cost-overruns, for unconscionable padding, and other fiscal abuses have been cavalierly ignored. The apotheosis accomplished, only reverence remains; the high priests, who could have elected to keep the emperor honestly naked, selected the path of least resistance and greatest personal return. Deserting the DOD for high places in civilian agencies of government, they perpetuated the mythology in unwonted places, with the result

[18] Ibid., p. 19.

[19] T. S. Kuhn, *The Structure of Scientific Revolution* (Chicago: University of Chicago Press, Third Impression, 1971). Kuhn calls the dominant paradigm a fundamental way of perceiving, thinking, and doing, consistent with a particular vision of reality.

that program budgeting (PPB)[20] became the iron law of administrative accounting; favorable cost/benefit ratios became the way to financial security; data, data-gathering, and information systems became an official obsession; and production and performance evaluation became the great national numbers game.

The literature on systems analysis in education burgeons with every conference on cost/effectiveness, with every symposium on managerial efficiency. "Experts," arising from economics, econometrics, operations research, engineering, sociology, education, and even philosophy are indistinguishable as to orientation. Quite contrary to early expectations that the systems approach would encourage multidisciplinary participation and creative synergism, experience reveals that almost irrespective of specialized input, the product comes out the same, a kind of amalgam in which any methodological rigor that had previously existed has been obliterated. Noteworthy here is the fact that a kind of law emerges, viz., the softer the discipline, the more amenable it is to the technological takeover. As might be expected as corollary to this, standards for expertness are practically nonexistent and accrue more to institutions than to individuals. Thus, persons who profess expertness in the management of solid waste systems claim without hesitation or qualification "systems competence" in education systems management and the Rand imprimatur virtually assures a favorable reception.

The weakness of relying on semantic similarity to transfer the techniques from the military and aerospace effort to social planning and education in particular has been discussed elsewhere.[21] The requisites for proper application of systems analysis have been specified but never satisfied.[22] This is not necessarily due to ignorance. Besides the caveats

---

[20]Program budgeting has been described as "a way of organizing cost data in such a manner that they can be used to analyze different courses of action in terms of cost and utility" (Lawrence Bogard, "Management in Institutions of Higher Learning," in *Papers on Efficiency in the Management of Higher Education*, by Alexander M. Mood, Cohn Bell, Helen Brownlee, Lawrence Bogard, and Joseph McCloskey (Berkeley, Calif.: The Carnegie Commission on Higher Education, 1972), p. 29.

[21]Ida R. Hoos, *Systems Analysis in Public Policy* (Berkeley, Calif.: University of California Press, 1973).

John S. Gilmore, John J. Regan, and William S. Gould, *Defense Systems Resources in the Civil Sector*, prepared for the U.S. Arms Control and Disarmament Agency (Washington, D.C.: U.S. Government Printing Office, 1967).

Anthony G. Oettinger and Sema Marks, "Educational Technology: New Myths and Old Realities," *Harvard Education Review*, 38 (Fall, 1968), p. 9.

[22]Robert Boguslaw, *The New Utopians* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1965).

A. M. Mood, "Cost-Benefit Analysis of Education," in *Analysis for Planning, Programming, Budgeting*, Mark Alfandary-Alexander (Ed.) (Potomac, Maryland: Washington Operations Research Council, 1968), pp. 17-26.

couched virtuously in footnotes in the plethora of papers offering input-output models, cost/effectiveness schema, and other mathematical exercises, there are forthright and professionally competent criticisms.[23] But the voices of reason are still and small compared with the clamor for "rationality," as paraded in the cloak of the techniques that will bring the "contributions from management science, operations research, organizational theory, and economics . . . to help solve administrative problems which typically confront colleges and universities" and, one might add, all other levels of education.[24] Mood has suggested that the first problem to be faced is whether the question to be studied is *suitable* for a system analysis.[25] Apparently, the new breed of "education managers" does not agree. Everything having to do with the educational process, from programs and policy at the topmost level to introducing a foreign language in the second grade in Chule Vista, California, is fitted into this Procrustean bed. Education planning has been forced into the same mold that was precast for getting the "biggest bang for the buck" (cost-effectiveness calculation in the Department of Defense) and manure-disposal in Manteca (Systems Analysis of Urban and Rural Waste, Aerojet General Corporation, 1968).

How does the "education manager" approach his task? Primarily, he is an information-gatherer. He is a model-builder. He "optimizes." Cognizant of financial stringency everywhere, he reasons that "improvement of operating efficiency" is the solution. "Efficiency," he says, "is measured as a ratio between two variables: cost and output."[26]

---

[23]As exemplified in writings by Guy Beneviste, "The Role of Utopian Models in Socioelectronics," invited paper, Research Committee on Sonotechnics, International Sociological Association, Loughborough, England, November 16-18, 1973; "Consequences of Excessive Educational Planning," paper presented at International Society of Educational Planners, "Science and Man in the Americas," American Association for the Advancement of Science, Mexico, June 25-30, 1973.

Paul Alper, "A Critical Appraisal of the Application of Systems Analysis to Educational Planning Models," *IEEE Transactions on Education*, E-11 (June, 1968), pp. 94-98.

F. E. Balderston, "Thinking about the Outputs of Higher Education," Paper P-5 (Berkeley, Calif.: Ford Foundation Program for Research in University Administration, May 1970).

Earl Cheit, "The Systems Challenge—How to be Academic though Systematic," paper delivered at annual meeting, American Council on Education, Washington, D.C., October 10, 1973. To be published in the Annual Proceedings.

[24]Paul W. Hamelman (Ed.), *Managing the University: A Systems Approach* (New York: Praeger Publishers, 1972).

[25]A. M. Mood, "Cost-Benefit Analysis of Education," in *Analysis for Planning, Programming, Budgeting*, Mark Alfandary-Alexander (Ed.), (Potomac Maryland: Washington Operations Research Council, 1968), p. 20.

[26]Howard R. Bowen and Gordon K. Douglass, *Efficiency in Liberal Education*, Carnegie Commission on Higher Education (McGraw-Hill Book Company, 1971), p. 3.

An increase in efficiency occurs, therefore, when output increases while cost remains constant. Education thus becomes a factory model—input-processor-output. University Hall, U.S.A., is busily at work generating management information systems, constructing models, and conducting cost/benefit comparisons in order to accomplish "analytic suboptimization," if nothing else.

These three activities merit scrutiny. The management information system (MIS) in higher education is

> a configuration of men, machines and methods which supports management in the collection, storage, processing, and transmission of information for operation, control, evaluation, and planning of a university. Seen in this way, the MIS draws upon the operating systems and supports both the PPB system and the administrative management and control system of the institution .... In an MIS recently designed for a state university the regular outputs to top administrators consist of 47 reports. Of these, 20 are planning reports and 27 concern control and evaluation of operations. The reports in their entirety cover the academic program students, faculty, support sciences, facilities, and finances.[37]

The generally conceded need for information has rationalized the most expensive, cumbersome, and demanding systems which, when tied into a computer, add confusion and complication while they consume precious resources. All the while that the information system is performing its sorceror's apprenticeship, simple bookkeeping operations are in disarray. At the University of California, for example, vital procedures, such as personnel and accounting, are almost hopeless bottlenecks, and the very information essential for proper model-building is nonexistent. The student-flow model is a case in point. Apparently, the objective here is to ascertain the "retention rate," itself a matter of dubious worth. Gathered for this purpose are figures on admissions, enrollments, and withdrawals. But when all is said and done, no one is the wiser, because no distinction has been made between students who drop out from the University of California to join the "drug culture" and those who transfer to Princeton; those who forsake formal education permanently and those who withdraw for a short period, perhaps even to carry on related field work. New and re-enrollments are undifferentiated, with the result that "retention rate" is calculated on an erroneous base, a hodge podge of duplication and double count.

[37]Charles A. Nelson, "Observations on the Scope of Higher Education Planning in the United States," Chapter 4 in *Managing the University: A Systems Approach*, Carnegie Commission on Higher Education (McGraw-Hill Book Company, 1971), pp. 36 and 37.

Forgotten in the frenetic numbers chase is the possibility that "retention rate" is not, in itself, cost/effective! It just may not be worth all the bother. Especially when considered in conjunction with the "relevant" curricula crafted to keep reluctant students in college, a high retention rate seems to be monumentally irrelevant to education.

The University of California "faculty flow model" has as inputs "economic" and "social" characteristics of the professorial staff. By whose judgment the latter are defined and by what reason the former are to be construed as the University's business are not clear. Since operation of the model is expected to yield "optimal hiring strategy," we can only conjecture about the analyst's criteria for "optimality."[28] What are his "significant social characteristics"—race, color, political affiliation, activism in causes? How will they be used in "plotting optimal hiring strategy?" Already, some of the effects of the University's planning-by-numbers are evident. Instead of valuing its older faculty members for the knowledge and wisdom that age presumably should have brought them, the University has been waging a vigorous campaign to encourage early retirement, so as "to provide planned opportunity for flexibility and continuing faculty renewal."[29] "Flexibility" could mean replacing senior faculty by temporary appointees; "faculty renewal" could deny tenure. Some optimum faculty workload model has dictated the elimination of half of the sabbatical leave program at the State Universities (San Francisco, San Jose, Hayward, et al.) and, by imposing restrictions on the nine campuses of the University of California, has had the effect of increasing the faculty's work load up to as much as one year in seven. The calculations have subtly demeaned the scholarly pursuits of professors, devalued the benefits accruing from the sabbatical programs, and given credence to figures which reflect inadequately and inaccurately how professors use their time.

Carrying further the "factory" syndrome, an optimum class scheduling model aims for more intensive "plant usage." Cost considerations and plant utilization, for example, influenced the decision to institute a quarterly, four-term calendar, possibly an economic success by some calculus, but, after four years of operation, still far from unequivocal as to who pays the costs and who derives the benefits. The possibility that compacting courses into 10-week sessions

[28] John Keller, "New Frontiers in Educational Program Planning and Budgeting and in Educational Cost/Benefit Analysis," speech at symposium at Berkeley, California, January 15, 1968.

[29] "The University Academic Plan, 1974–1978," draft released by Executive Vice President Chester O. McCorkle, October, 1973.

or, as is now contemplated, scheduling evening classes might interfere with employment or other activities of the student's life or adversely affect the teaching-learning process does not appear in any model, *nor* do classrooms with standing-room-only conditions while a 900-seat auditorium remains unused, because "all those empty seats would have affected the utilization ratio that the state calculates for classroom usage."[30]

The "favorable ratio" phenomenon is not unique to the University of California, discussed here because this institution *is* typical and also somewhat in the vanguard as to utilization of these "managerial concepts." Not unexpectedly, the imposition of the techniques has moved far out of fiscal affairs and into matters academic. "Plant utilization" controls course offerings; "efficiency studies" determine the fate of departments; "performance" and "productivity" are serious criteria for evaluation. When, as now seems to be the case, number of degrees awarded is a basis for fund allocation, an interesting feedback effect becomes obvious. By easing its requirements and lowering its standards, a department could improve its financial position even though, at the same time, it could run the risk of losing its accreditation! As for the students, caught in the speed-up processing by the "knowledge factory," they will emerge from the output side of the model with deliberately devalued paper that is called a diploma or a degree.

Performance has become so pervasive a criterion that at the college level professors are being rated on this basis. Criteria are, of course, scanty and slippery, for beyond number of hours of teaching and number of journal articles written, there are only the elusive qualitative factors. Recommendations by an ad hoc review committee at Berkeley point up the new trend—evaluations of professors' performance by students. Besides constituting an abdication of the responsibility of maintaining academic excellence through properly authorized professional channels, this kind of procedure allows for the most flagrant abuse, no distinction being made between the rating by those who earned A's and those who failed the course. If professors are forced to pander to popular preferences in order to survive, we may yet see a paid claque performing in the classroom!

At the elementary and secondary school levels, emphasis on "performance" and "productivity" have taken on somewhat different, but nonetheless portentous, manifestations and implications. Ac-

[30]Statement of Jean Dobrzensky, of Chancellor's Office, University of California, as quoted in *Daily Californian*, January 24, 1974.

companying the predilection to apply performance measures to school programs is the necessity to cast pupils in the role of producers whose "output" represents the basis for calculation of cost/effectiveness ratios. Even though PPBS as such is no longer obligatory, its official demise occurred after the point of no return had been passed. Its vestigial remains hang on under new names, and administrators at state, district, and local levels will be implementing program budgeting in some fashion for some time to come, because they know neither what to do nor how to undo what has already been done with it. More vulnerable than ever to the hard-sales campaign of curriculum peddlers, they have made such a fetish of "performance" that children have been cast into the role of "performers," to the detriment of other, perhaps more important, phases of development. In keeping with this emphasis, "behavior modification" has been revived as a legitimate activity;[31] "chemical straitjackets," in the form of amphetamines, have been used to induce docility and conformity in the classroom and thus assure favorable "output" for teachers, programs or whatever is being weighed in the cost/effective balance. Skinnerism and other forms of latter-day behaviorism are gaining followers in psychology, sociology, and education.

Another direct outcome of the pressure for performance is the practice of letting contracts to private companies to conduct specified courses in the public schools. In many cases, compensation is supposed to be made according to the actual results achieved by the students. Funded by Title I of the Elementary and Secondary Education Act of 1965, contracts had, by the end of 1970, been let by 18 school districts in 16 states, as part of the Office of Economic Opportunity program, to six companies: Alpha Learning Systems, Learning Foundations, Plan Education Centers, Singer Graflex, Westinghouse Learning, and Quality Educational Development. "Management support services" were supplied to the entire project by Education Turnkey Systems. An industry enthusiast, finds the "influx of systems people and companies into the staid area of education particularly exciting" and claims that the unorthodox methods, such as rewarding children with candy, extra free time, or trading stamps, must be good because the National Education Association has attacked the concept in Congress, the American Federation of Teachers has declared the idea dehumanizing,

[31]Subcommittee of the Committee on Government Operations, House of Representatives, Ninety-first Congress, Second Session, "Federal Involvement in the Use of Behavior Modification Drugs on Grammar School Children of the Right to Privacy Inquiry," September 29, 1970 (Washington, D.C.: U.S. Government Printing Office, 1970.

and the United Federation of Teachers has sought a court injunction to prevent continuation![32]

Evaluation of the OEO experiment, which included some 27,000 students and cost about $6 million, was, fortunately, not left to interested companies, instances of orienting curriculum toward the tests and other such abuses having already been revealed. The Comptroller General's Report to the Congress provided a noteworthy overview, summarized in the following paragraph:

> Was performance contracting more successful than traditional classroom instruction in improving the reading and mathematics skills of poor children? The answer according to OEO is no! OEO's report released in June 1972 stated that "the results of the experiment clearly indicate that the firms operating under performance contracts did not perform significantly better than the more traditional school systems.[33]

The imposition upon teachers of performance measures stems from the pervasive management-efficiency syndrome and stretches from coast to coast. In New York State, a performance-based certification is being imposed on teachers; in California, the Stull Act requires that teachers be judged on the basis of their pupils' achievement. The New York plan has been called an attempt "to evaluate and certify teachers on a 'piecework' basis, i.e., how many reading score points were made per child."[34] According to official literature, the Stull Act (Assembly Bill 293, July 20, 1971) "provides a unique opportunity for school boards and certificated staff members to respond to parent, citizen, and legislative demand for educational accountability and to improve instructional programs."[35]

Especially enlightening is Item C, *Criteria for Evaluation:*

> Article 5.5, Section 13487 of the Stull Act has identified four specific steps in certificated personnel evaluation:
> (a) The establishment of standards of expected student progress in each area of study and techniques for the assessment of that progress.
> (b) Assessment of certificated personnel competence as it relates to the established standards.

[32]Thomas DeMarco (Vice President, Mandate Systems, Inc., New York), "Software for Schools," *Modern Data*, (December, 1970), p. 46.

[33]Comptroller General of the United States, Report to Congress, "Evaluation of the Office of Economic Opportunity's Performance Contracting Experiment," B-130515, May 8, 1973, pp. 1-2.

[34]Sandra Feldman, "Performance-Based Certification: A Teacher Unionist's View," presented at AERA, New Orleans, n.d., pp. 5-6.

[35]California State Department of Education," Certificated Personnel Evaluation in California Public Schools. Implementing the Stull Act," Sacramento, California 1973, p. 1.

(c) Assessment of other duties normally required to be performed by certificated employees as an adjunct to their regular assignments.
(d) The establishment of procedures and techniques for ascertaining that the certificated employee is maintaining proper control and is preserving a suitable environment.[36]

Underlying the verbal macramé of the Act and its attendant documents is a shotgun approach to performance. Section (a) in the above list of particulars demands that teachers, in order to maintain their certification, must guarantee students' progress "in each area of study." This being a psychological and pedagogical impossibility, the Stull Act makes liars out of teachers or teachers out of the best liars. Just as the most egregiously brazen of the performance contractors manufactured the most convincing and self-serving "evidence," so will those teachers with the fewest scruples survive the trial by pupil progress and performance.

CONCLUSION

The assumption that education is a system and, therefore, subject to the quantified techniques developed for weaponry systems, used in space exploration and applied ubiquitously, has brought it to a perilous state. Notions of industry-style efficiency have put a premium on limited, and not necessarily the most important, aspects of education. Where humane and individual approaches are called for, paper-and-pencil games ordain mass "solutions." Preoccupation with numbers has subverted educational theory and philosophy. Accountability substitutes for integrity; the final accounting is a printout that, like an infinite hall of mirrors of self-reflection, creates only an endless regress. The cause of efficiency may thus be served, but the costs are incalculable and will have to be borne by the whole society, now and in the future.

[36] Ibid., p. 5.

Journal of Dynamic Systems, Measurement and Control

forum

A continuing feature in which noted individuals in various areas are solicited to report on what is and should be going on in their fields. Response is welcomed and will appear in the Discussion section of future issues.

Technology Transfer: Another Opinion

K. PRESTON WHITE, JR.,[1] AND DONALD WRIGHT.[2] In a recent Forum [1],[3] Ida Hoos proposed to argue the case against "Systems Analysis as Technology Transfer." The diverse issues broached in the Hoos essay are undeniably appropriate subjects for serious debate in this *Journal*. These include the compass of the systems approach, the viability of analytic planning, and the social responsibilities of the engineering community. The Hoos essay, however, is a distressing embarkation on such debate.

Given a narrow construction, certain elements of the Hoos argument are not entirely without merit. Assuredly, some "systems analyses" of social problems have failed. And the reason for these failures may be traced, in part, to particular analysists' lack of understanding of the subject they were treating. Indeed, quantitative systems analysis *as commonly practiced in engineering and management* is not directly applicable to systems which are highly complex or varied, or which are incapable of precise definition or description, or for which there exist no entirely acceptable or optimal solutions. Many of our most pressing social problems are rooted in systems which are characterized by all of these difficulties. Therefore, the engineering brand of systems analysis is not applicable to social systems. Engineers whose backgrounds are limited to experiences with engineering and management systems are not competent social planners.

In all of this, we concur. The engineering community—and the public, in general—is well disabused of notions that a technology can single-handedly remedy all the ills of our society. In fairness, however, much the same must be said for the techniques and perspectives of any individual professional community. Our major societal problems are neither solely technological, nor sociological, nor psychological, nor political, nor economic, nor scientific. Our problems are all of these, together, and more. It would seem that professional communities must work in concert if there is to be hope for major societal improvements. Perhaps one of the most immediate issues confronting the engineering community is the need for introspection and self-criticism, to the end that the community may discern its potential contributions to societal improvement and its consequent role in interdisciplinary efforts.

Our specific contentions with the Hoos essay are twofold. First, readers of this Forum unfamiliar with the extent of the family of disciplines based upon systems theory are left with the impression that all systems constructs are useless outside of engineering applications. Second, readers of this Forum unfamiliar with the theses of Thomas Kuhn [2] (upon which the Hoos essay heavily relies) are left entangled in a web of positive and normative assertions regarding the engineering community and the "trained incapacity" of engineers to engage in social planning. It is to correct such misimpressions—or at least to present another side of a case which legitimately may have more than one side—that the present essay is directed.

Systems Disciplines

The *systems view* is a philosophical perspective which asserts universal holism, that is, asserts that every entity in the universe is potentially related, to a greater or a lesser degree, to every other entity. According to the quality and quantity of these relationships, entities may be classified together as forming systems. In its most general sense, the *systems approach* is a method of inquiry which adopts and utilizes the systems view.

Systems theory is a science which attempts to formalize, in terms of logic and mathematics, the systems view. A component of this formalization distinguishes two basic properties of every system: (1) its behavior, the pattern relationships between the system and its environment; and (2) its structure, "the manner of arrangement (organization) of the mutual couplings between the elements of the system [and] the behavior of these elements" [3]. Systems theory also distinguishes basic types of problems concerning systems, among these: (1) *systems synthesis*, the problem of delimiting a system's structure from a known behavior; and (2) *systems analysis*, the problem of determining a system's behavior from a known structure.

Contrary to the impression fostered by the Hoos' essay, systems theory is generally agreed to have originated in biological science, not in engineering. The early development of systems thought is associated with the biologist Ludwig von Bertanlanffy, especially that of General System Theory [4]:

"From the statements we have made, a stupendous perspective emerges, a vista towards a hitherto unsuspected unity of the conception of the world. Similar general principles have evolved everywhere, whether we are dealing with inanimate things, organisms, mental or social pressures. What is the origin of these correspondences?

"We answer this question by the claim for a new realm of science, which we call General System Theory. It is a logicomathematical field, the subject matter of which is the formulation and derivation of those principles which hold for systems in general. A 'system' can be defined as a complex of elements standing in interaction. There are general principles holding for systems, irrespective of the nature of the component elements and of the relations or forces between them. From the fact that all the fields mentioned are sciences concerned with systems follows the structural conformity or 'logical homology' of laws in different realms.

"The principles that hold for systems in general can be defined in mathematical language. It will be seen then the notions such as wholeness and sum, progressive mechanization, centralization, leading parts, hierarchical order, individuality, finality, equifinality, etc., can be derived from a general definition of systems: notions that hitherto have often been conceived in a vague, anthromorphic, or metaphysical way, but actually are consequences of formal characteristics of systems, or of certain system conditions."

[1]Graduate Student, Department of Mechanical Engineering and Materials Science, Duke University, Durham, N.C.

[2]Associate Professor, Department of Mechanical Engineering and Materials Science, Duke University, Durham, N.C.

[3]Numbers in brackets designate References at end of Forum.

General Systems Theory, then, is vitally concerned with systems on the general level at which C. West Churchman defines them [5]: "a set of interrelated components working together to accomplish a certain task." A popular (though mathematically nonrigorous) application of general systems conceptioning to societal problems is found in R. Buckminster Fuller's *Operating Manual for Spaceship Earth* [6].

The major subdisciplines of systems theory are concerned with different realizations of relationships among the elements of general systems. The natural sciences have long studied systems from the perspective of mass-energy relationships. A relatively new science, *Cybernetics* (see, notably, Weiner [7] and Beer [8]), deals with general systems from the perspectives of organization, communication, and control. Cybernetics is strongly tied with a third subdiscipline, *information theory*, the science of information and communication within systems. Again, these disciplines are equipped to deal with systems on a general level. As another reader of this Forum has suggested [9], with regard to physician W. Ross Ashbey's *An Introduction to Cybernetics* [10], cybernetics also may have a significant practical value in engineering. The practical value of mass-energy and information theories is surely undoubted.

Other systems disciplines are less general and are biased toward consideration of one or another specific class of systems. In engineering, *control theory*, *information processing systems*, and *communications systems* are examples of cybernetic subspecialties. In urban and regional planning, a discipline increasingly called *systemic planning* is another. The works of Chadwick [4], McLoughlin [11], and Catanese and Steiss [12], are notable in this context.

Finally, certain systems disciplines are highly specialized. These disciplines employ methods and techniques which are predicated upon fundamental (and too often implicit) assumptions regarding the systems question. Among these disciplines are, for example, *linear control systems*, *operations research*, and *systems engineering*. Systems analysis, the undefined discipline to which the Hoos essay is directed, may be thought to comprise both operations research and systems engineering. The assumptions, methods, and tools of systems analysis need not be catalogued here—it suffices to say they are not always well-suited to the consideration of complex social issues. This point has been made before and often, most forcefully by Catanese and Steiss [11]. The point should be obvious by now.

Mistaken Impressions

In attempting to reiterate this simple point, the Hoos essay fails to delimit with coherence the problem which it addresses. The result of this oversight is the creation of at least two mistaken impressions. First, the essay implies that none of the more general systems disciplines are of any potential use in comprehensive planning. The implication is at least debatable and we refer the reader to the previously-cited works on planning for a counter-impression. Second, the essay fosters the impression that management science and systems engineering invariably fail, even when applied to more narrowly-defined problems in societal planning. This implication is not debatable—it is simply untrue.

A lengthy quote from a review of Dr. Hoos' book, *Systems Analysis in Public Policy* (from which the Forum essay undoubtedly is abstracted), is in order. Reviewer Richard de Neufville cites many of our own objections [13]:

"The text itself appears disjointed and fails to build a compelling argument. The repeated examples of poor analysis, which constitute the main part of the book, do not demonstrate whether all systems analysis is bad or only part of it is. There is not even any indication of whether the examples cited are representative of actual applications. As it happens, they frequently are not. For example, under the rubric 'the techniques at work in waste management,' she describes a single case. While this one application may be inept, it is quite atypical of how systems analysis actually is being used to expedite garbage collection (with some noticeable success, I might add). Similarly, she repeatedly and unjustifiably ascribes particular statements to the whole field. She would, for example, have the reader believe that just about all systems analysts were imbued with the idea that since we have been able to get man on the moon we can solve any problem. Particular anecdotes do not describe a problem, and they certainly leave the reader with no guidelines for judging the merits of any particular analysis.

"The presentation is not only weak in concept but also in detail. It is inaccurate and superficial in form.... The author's knowledge of systems techniques is, from the context, quite weak. Her discussions of areas with which I am particularly familiar, social indicators for instance, repeatedly draws inferences which are not, by any reasonable test that I know of, supported by the evidence presented. Furthermore, she is quick to impugn motives and to accept slanderous comments uncritically. My overall impression is that the text was put together carelessly...

"The author has really not given us a reasoned critique, but a diatribe. My impression is that she started with the opinion that analysis was not only wasteful, but insidious and evil as well. It would then seem that she had picked upon systems analysis as a particularly vulnerable aspect of analytic planning, and launched into an attack without seeing the necessity to build a careful case. I suspect that the text should be seen as a moralistic statement against the rational model of the planning process. As such, it would be part of the growing antitechnology movement that is evident in the press and on campuses."

Semantics and Semantics

Much of the confusion of the Hoos essay results from the "slippery semantic" issues, about which that essay makes much ado. The Hoos essay identifies two "booby traps" which allegedly are responsible for unwarranted attempts to apply systems analysis to social systems. The first booby trap is "the loose lexicology of the word, *system*." To prove that the concept is vague, the reader is referred to the seventeen definitions provided in Webster's dictionary. The reader is assured there are more and is treated to the intemporate assertion that, "Every writer on the subject seeks to achieve refinement of conceptualization through verbalization, a not uncommon ploy of the written word."

Ploy? We must make the following observations. First, Noah Webster is rarely an adequate reference for technical definitions (such has been our repeated experience). Second, although the technical definitions of "system" are indeed manifold, ranging from the verbal to the mathematical, they are each consistent with the other (none of the technical definitions to which we have been exposed in any way contradict that of Churchman—which even the Hoos essay concedes comprises "the generally accepted ingredients of a working definition.") Third, if the reader requires a rigorous logico-mathematical definition, he is referred to Ellis and Lugwig [14]. Finally, we suggest that the definition of system is not vague, i.e., imprecisely defined; rather it is general, i.e., of wide applicability.

The issue of lexicology is not entirely without foundation, however. If we concede that scientists—natural, applied, and social—can find meaning in the general concept of a system, where are the "slippery semantics"? They are, we believe, in the definitions of the *ménage* of theory, method, and technique abutted with the prefix "systems." "Systems analysis," "systems capability," "systems methods," "systems expertise," "systems approach," "systems engineering"—all of these are used with seeming interchangeability in the Hoos essay. The casual reader, unfamiliar with the family of systems disciplines, is left with the impression that these and all other systems constructs mean quite the same thing. They do not.

The second booby trap, is, in Hoos' words, "fallacious analogy, for to impute verisimilitude where only superficial similarity exists leads to nonsensical conclusions about everything, including skill transfer." This trap is attributed to the "vagueness"—"generality," if you permit—of the word, system. This observation is close to the mark, but the problem does not rest entirely in the generality of the concept of a system. The burden of the problem lies with the concepts of "analogy," "verisimilitude," and "similarity." It is the misunderstanding of these concepts—concepts which are so much the domain and concern of a higher order of "systems expertise"—which in some instances has led to nonsensical conclusions about skill transfer.

The Hoos essay is concerned that the readers of this Forum have a faulty understanding of what constitutes "systems capability." If this concern is valid, then the essay is remiss in simply expounding upon what a systems capability is not. The uninformed reader is left with the impression that such capability simply is not. And this is untrue.

Think of an analogous situation in engineering education. A dynamics professor is concerned that his students have a faulty understanding of Newton's law. Students have been applying $f = ma$ to problems of variable mass. As a consequence, they have been wasting their time, burning up valuable resources in graphite and theme paper, and the problems are still unsolved. To correct this "amorphous state of conceptualization," the professor lectures that it has been repeatedly demonstrated that $f = ma$ does not apply to rocket problems, because the mass of the rocket varies. What has this lesson taught? Hopefully, it has taught that *$f = ma$ does not apply to variable-mass problems.* For students that know Newton's law only as $f = ma$, the lesson also has taught that *Newton's law does not apply to rocket problems,* which clearly is untrue and is not what the professor intended. Has the students' conceptualization if rocket problems or Newton's law been improved? Yes and no. What the Hoos essay teaches is that systems analysis does not apply to social problems. To the reader who knows the field of "systems" only as systems analysis, the lesson has also taught that a "systems approach" does not apply to social problems. The worth of this lesson we leave to the reader's judgment.

Paradigms

Our second contention with the Hoos essay regards its characterization of the engineering community. This characterization leads to suggestions as to the appropriate means by which individual engineers and the engineering community at large must seek social relevance. To appreciate fully what the Hoos essay attempts to do, one must be familar with the theses Thomas Kuhn advances in his exceptional essay, *The Structure of Scientific Revolutions* [2]. To do justice to Kuhn's essay in this limited space, of course, is impossible. Therefore we shall attempt to summarize only two theses which are particularly germane to the present discussion.

Kuhn invokes two related concepts, both of which are meant by the word *paradigm.* The first and general meaning of paradigm is, in Kuhn's words, "disciplinary matrix": examples, beliefs, models, values, generalities, and so on, shared without dissent by members of a professional community. It is this sense of the word which the Hoos essay intends in its definition of paradigm "as a fundamental way of perceiving, thinking, and doing, consistent with a particular vision of reality."

The second, more specific, and more original meaning of paradigm is, again in Kuhn's words, "exemplars": shared examples which the student experiences as an essential part of his initiation into the particular vision of reality held by the professional community. Kuhn's thesis is that the science student assimilates "a time-tested and group-licensed way of seeing" by "doing science rather than by acquiring rules of doing it." Solving problems—both textbook and laboratory problems—does not merely or even primarily serve as a medium for gaining facility in applying the theory and rules of science, rather "doing problems is learning consequential things about reality" [2]:

> "The student discovers, with or without the assistance of his instructor, a way to see his problem as *like* a problem he has already encountered. Having seen the resemblance, grasped the analogy between two or more distinct problems, he can interrelate symbols and attach them to nature in the ways that have proved effective before. The law-sketch, say $f = ma$, has functioned as a tool, informing the student what similarities to look for, signaling the gestalt in which the situation is to be seen. The resultant ability to see a variety of situations as like each other, as subjects for $f = ma$, or some other symbolic generalization, is, I think, the main thing a student acquires by doing exemplary problems, whether with a pencil and paper or in a well-designed laboratory. After he has completed a certain number, which may vary widely from one individual to the next, he views the situations that confront him as a scientist in the same gestalt as other members of his specialists' group. For him they are no longer the same situations he has encountered when his training began."

Engineering educators, no doubt, already are convinced of this thesis.

Perhaps Hoos' strongest argument opposing engineers' endeavors in social planning is founded upon these concepts of paradigms. The exemplars which the engineering student learns are not always consonant with exemplars relevant to social systems. The vision of reality the engineering student acquires through the assimilation of these exemplars—the vision of reality of the engineering community—is not always a vision appropriate to social planning. The engineer "sees" systems as "things;" social systems are "ideas." The engineer tackles a problem as though it had a single, objective "solution;" social problems have no consensus "solutions," *per se.* Engineering problems are "value-free;" social problems are "value-laden."

As a general assessment of the majority engineering community, this recognition is compelling. However, it is no more than a generality. The engineering community is not so monolithic as the Hoos essay would have us believe it is, or should be, or must be. There are substantial numbers of engineers whose vision of reality *is* consonant with the reality of social systems. This vision has been conditioned not only upon "traditional" engineering exemplars, but also upon exemplars of the social sciences and of general systems theory. This "different vision" has not precluded their competence as practicing engineers. It has allowed them competence in planning.

There is, in our opinion, a particular aspect of the engineering general paradigm which, when coupled with the appropriate exemplars, allows engineers to be particularly good planners. This is, for want of a better phrase, "purposeful, explicit assumption-making." The engineer is willing to make assumptions about reality, not because he believes that his assumptions are true (in some objective sense), but because they "work." "What works," in turn, only has meaning in terms of one's goals and objectives. A digression upon pragmatic philosophy, linguists, or gestalt psychology is beyond our competence (our rather limited introduction to these subjects includes Dewey [15], Russel [16], Polanyi [17], and Postman and Weingartner [18]), but we suggest that the engineering approach creates a "focal awareness" (to borrow Polanyi's term) of the process of assumption-making. In engineering, this awareness bounds the process of finding solutions. Coupled with the appropriate exemplars and a cautious redefinition of "what works," this awareness can be the driving force behind a planning process which seeks not "social solutions," but "social improvement." Certainly improvement may be possible.

The Engineering Community—Positive Assertions

A second of Kuhn's theses is that, to understand science, it is necessary to study the structure of the scientific community. Although Kuhn refers expressly to science, not technology, the thesis extends to include other disciplines as well.

"I shall close by underscoring the need for similar and, above all, for comparative study of the corresponding communities in other fields."

Kuhn proposes a number of questions appropriate to this study.

"How does one elect and how is one elected to membership in a particular community, scientific or not? What is the process and what are the stages of socialization to the group? What does the group collectively see as its goals; what deviations, individual or collective, will it tolerate; and how does it control impermissible abberation?"

These are precisely the questions which, for the engineering community, the Hoos essay poses answers. These answers are somewhat confused between the positive and normative modes—between the "is" and the "ought"—but they are worthy of unravelling.

In the positive mode, the Hoos essay asserts that engineers elect to membership because they have "a basic lack of sociability" and "low people-orientation." The socialization process begins with exemplars that are "thing-" and not "people-" or "idea-oriented." The engineer is isolated physically and ethically from the society at large. The resulting ethical myopia is reinforced by his station as an employee of industry or the military: he has little control over the technological change he helps to create and consequently refuses to accept responsibility for that change. He learns to do what he is told. He seeks solice in "doing his thing well," the principal value of the group. The community also values ethical conduct between engineer and colleague and between engineer and employer, but not between engineer and society at large. The values of the group are enforced, presumably, by the marketplace.

The Hoos essay also recognizes a new and dissonant element in the socialization process. At some time in the process of being or becoming an engineer, one is afflicted with "what the psychiatrists call a 'collective unconscious guilt'." This guilt is a product of the engineering community's "heritage couched in the military." It is to assuage this guilt that the engineer seeks "relevance." The engineering community has come not only to condone relevance-seeking as appropriate activity for engineers, but has come to adopt relevance itself as a group value. "Engineers *en masse* are exhorted by deans of their colleges to turn their talents to solving the amorphous mess known as social problems..."

Following Kuhn's work, this dissonance is the beginning of a guilt-induced "paradigm-shift." The engineering community is redefining its values, changing its self-image and the image of what engineering is about. Hoos desires to forstall such a shift. The problem, as Hoos sees it, is that the tools and exemplars of the community are not changing accordingly. Relevance is being sought in the wrong place—social planning—where engineers have no competence.

Before considering the normative assertions of the Hoos essay, it is appropriate that we look at this posivive characterization more closely. As a *caricature*, this description of the community may not be far from fact. Unpleasant, but with elements of truth. It is, of course, no more than a caricature—a generality with salient features exaggerated.

What we find difficult to accept is the notion of "collective unconscious guilt" as the motivating force behind relevance-seeking. We earnestly wish to question Hoos' sources to this point, but she cites none (which, in light of her scrupulous documentation of Shakespeare, is particularly odd). "With no pejorative intent," we suspicion that Dr. Hoos' knowledge of psychiatric science is on a par with her understanding of systems science. At any rate, there seems an equally plausible explanation. This is the very *conscious* recognition that, here and now, the engineering community is participating in activities which are, to some, morally repugnant. Seeking relevance is seeking alternatives to participation in those activities, in real time.

We have steadfastly endeavored to refrain from the emotionalism which characterizes the Hoos essay, but there is one image which we can not seem to shake. If the reader will please forbear, we are reminded of Arlo Guthrie's reaction, in *Alice's Restaurant* [16], when he finds he has been withdrawn from the selective service pool, because he was once convicted of littering in Stockbridge:

"I mean...I'm sittin' here on the Group-W bench, because you want to know if I'm moral enough to join the Army, burn women, kids, houses and villages, after being a litterbug?"

We have often felt *that* way about engineering.

The Engineering Community—Normative Assertions

The only consistent, normative motif of the Hoos essay, regarding what the engineering community should do, appears to be the prejudice that engineers should avoid social planning, because they are and always will be inept at it. Refusing to admit any present competence, or any possibility of future competence, the Hoos essay advises that the community discourage "crossover" or "passing," not only among engineering disciplines, but between engineering and the social sciences. It is suggested that such discouragement can be affected by demanding specialization: a long history of specialization has made engineering great and justifies a return to the original paradigm.

Specialization, of course, *will* keep engineers out of the social arena. But the problem of "guilt" (or current irrelevance of the old paradigm) is not solved. The dissonance giving rise to the new paradigm cannot be wished away so easily. History provides no justification, in itself, for we are told (in almost the next breath) that it is this same history which causes the "guilt."

If he (or she), must purge himself of this guilt, the engineer is advised that there are many relevant activities which he should pursue within his specialty. An admirable list is offered. But if we are operating under the old paradigm of engineer-as-employee, doing well only that which he is told to do, from where does the power emanate by which the engineer demands the opportunity to pursue these relevant activities? Employers, including the manufacturers of medical equipment, construction companies, and Detroit, are not necessarily motivated by an altruistic concern for safer products. Profit is often an overriding concern. Either relevant projects appear without the intervention of the engineering community (industry changes its paradigm), or the engineering community must learn to demand their undertakings. Specialization is not the key to that kind of learning.

The second, guilt-relieving suggestion offered by the Hoos essay is denegation.

"Perhaps it is not the engineers who should be blamed for behaving true to form, in only doing what is asked of them... Perhaps we should look farther and direct our criticism to government agencies who should know better."

Two questions surface in this regard. First, can this strategy be successful within the old paradigm? Second, if it can, should the community adopt this strategy? To the first and positive question, we submit that simple denial is precisely the strategy which historically has led to guilt. Denial, coupled with rational analysis, however, might work. The level of consciousness of the group could be raised and consciousness of the previously unconscious guilt would allow its treatment. The community could then accept its employee status and no longer feel guilt in regard to that over which it has no control. But this new consciousness is of the new paradigm, not the old. And it permits a positive rephrasing of the second and normative question. Would such an enlightened community actually have no control over the results of technology?

We think that it would. The traditional "cop-out" of the engineer has been, "if I didn't do it, someone else would." This rationalization would no longer avail in an enlightened community. That "someone else" is invariably an engineer. The enlightened engineering community would militate against the "someone elses" among its fold.

Regardless of the viability of Hoos' prejudice, the old paradigm is no longer viable once the engineering community acknowledges that it has made and is making ethical mistakes with regard to society at large. The paradigm has changed and there is no returning. Specialization can not force such a return. The options open to the community are either to do something assertive to gain relevance, or learn to live with irrelevance. The latter course, we feel, is highly undesirable.

Conclusions

In the first part of this essay, we have attempted to forestall the equating of "systems analysis" with the entire field of systems disciplines. With this erroneous equation goes the notion that a systems approach is valueless outside of engineering. We have suggested, in fact, that the more general systems disciplines can and have had some success in social planning, albeit at a strategic level. We have offered the opinion that an expertise in these general constructs, coupled with an understanding of social phenomena, eventually may lead to the development of specific tools for social systems analysis.

In the second part of this essay, we have attempted to show that specialization is not a curative for the engineering community's dilemma of social responsibility. The mode of "doing well what one is told to do" will be dissonant with the value of "achieving something relevant" as long as the engineering community remains the irresponsible servant of other social communities. Specialization can help the engineering community to do what it does well, but can not insure the relevance of that activity. It is encumbant upon the engineering community to ask the question, "What is worth doing?" It is encumbant upon the engineering community to seek the means by which they may act upon the answers to this question.

The present trend in engineering education is of the new paradigm of awareness and social responsibility. This paradigm may provide a resolution to the community's dilemma: a socialization process which stresses the community's ethical and moral responsibilities to society at large. If the trend persists, the engineer eventually may be able to say, "If I don't, *nobody* else will."

Aspects of the problem identified in the Hoos essay are valid, but the solution is not simply to retreat into specialization. The engineer wishing a specialty should have one. The community should guarantee that the activities of that specialty are relevant. But the engineer wishing to tackle problems of generality should not be dissuaded. The community should encourage him and provide him with the appropriate theory and exemplars. It is our opinion that the theory of general systems and the combined exemplars of engineering and social science are appropriate to that task. It is the engineering generalist who will create a role for himself in much needed interdisciplinary efforts.

The major problems of our society may have no immediate or enduring solutions. If such solutions do exist, they will never be found through the techniques of any particular discipline, or any combination of such techniques. The tools appropriate to that quest are trust, generosity, and human understanding. But the lack of solutions should not end the search for improvement. No doubt, there is much that can be improved.

Acknowledgments

The authors greatly appreciate the thoughtful criticisms and suggestions made by our graduate student colleagues and especially by Dr. N. B. Ratliff, Jr. The responsibility for the opinions expressed in this essay, of course, remains our own.

References

1 Hoos, I. R., "Systems Analysis as Technology Transfer," JOURNAL OF DYNAMIC SYSTEMS, MEASUREMENT, AND CONTROL, TRANS. ASME, Vol. 96, No. 1, Mar. 1974, pp. 1–5.
2 Kuhn, T. S., *The Structure of Scientific Revolutions*, University of Chicago Press, Second Edition, Fourth Impression, 1973.
3 Klir, Jiri, and Valach, Miroslav, *Cybernetic Modelling*, D. Van Nostrand Co., Inc., 1966.
4 Chadwick, George, *A Systems View of Planning*, Pergamon Press, 1970.
5 Churchman, C. W., quoted in Hoos, *op. cit.*, p. 1.
6 Fuller, R. B., *Operating Manual for Spaceship Earth*, Pocket Book Edition, Fourth Printing, 1972.
7 Weiner, Norbert, *Cybernetics*, Wiley, 1948.
8 Beer, Stafford, *Cybernetics and Management*, Wiley, 1952.
9 Porter, B., "Cybernetics and Control Engineering," JOURNAL OF DYNAMIC SYSTEMS, MEASUREMENT, AND CONTROL, TRANS. ASME, Vol. 95, No. 4, Dec. 1973, p. 349.
10 Ashbey, R. W., *An Introduction to Cybernetics*, Wiley, Third Impression, 1958.
11 McLoughlin, J. B., *Urban and Regional Planning: A Systems Approach*, Praeger Publishers, 1969.
12 Catanese, A. J., and Steiss, A. W., *Systemic Planning: Theory and Application*, Lexington Books, Second Printing, 1972.
13 de Neufville, Richard, "Book Reviews: *Systems Analysis in Public Policy. A Critique*—Ida R. Hoos," *Transactions on Systems, Man and Cybernetics*, IEEE, Sept. 1973, pp. 532–533.
14 Ellis, D. O., and Ludwig, F. J., *Systems Philosophy*, Prentice-Hall, Inc. 1962.
15 Dewey, John, *The Quest for Certainty*, Capricorn Books Edition, Tenth Impression, 1960.
16 Russel, Bertrand, *The Impact of Science on Society*, Simon and Schuster, 1953.
17 Polanyi, Michael, *Personal Knowledge*, Harper Torchbook Edition, 1964.
18 Postman, Neil, and Weingartner, Charles, *Teaching as a Subversive Activity*, Delacorte Press, Third Printing, 1969.
19 Guthrie, Arlo, *Alice's Restaurant*, Reprise Records.

Author's Reply to Technology Transfer: Another Opinion[1]

White and Wright's "other opinion" is to be commended, especially for reinforcing, sometimes positively, more often negatively, many of the salient points made in my *Forum* article (March, 1974). One might have wished, however, that they had not fallen into several of the booby traps I warned against. Their main concerns are that the reader was left (*a*) "with the impression that all systems constructs are useless outside of engineering applications," and (*b*) "entangled in a web of positive and normative assertions regarding the engineering community."

The first item illustrates tellingly the very semantic slippage which I described and which they criticize me for emphasizing. Their scholarly dissertation on lexicology sheds scant elucidation. In it, they make the usual appeal for faith in the meaning implicit in "the general concept of a system" but fail to articulate what this meaning is. Then, they perform a delicate *pas de deux*

[1]By K. Preston White and Donald Wright.

about the perceptions of the discerning, as opposed to those of the casual, reader. With the claim that I misled the latter into thinking that all sorts of systems constructs mean the same thing, they state authoritatively and without definition, "They do not." This kind of disputation is straight out of *Alice in Wonderland*, her famous dialogue with Humpty Dumpty going as follows:

> "When I use a word," Humpty Dumpty said in a rather scornful tone, "it means just what I choose it to mean—neither more nor less."
>
> "The question is," said Alice, "whether you can make words mean so many different things."
>
> "The question is," said Humpty Dumpty, "which is to be master—that's all."[2]

Humpty Dumpty makes an important point; the *Weltanschauung* (world view) of the analyst, his perception of the system, his selection of variables as to relevance and pertinence, his assignment of priorities, his value system—all of these are crucial in dealing with the spectrum of constructs we have crudely labelled *systems*, and the more so when dealing with the amorphous construct, *social systems*. In short, the vital factor is *who is master*, that's all.

Humpty Dumpty thus brings us to White and Wright's second bone of contention, namely, the *engineer* as system designer, architect, or analyzer. In choosing to focus on my discussion of the trained incapacity of engineers, they inadvertently substantiate rather than refute my position. They could have argued that trained incapacity, even though originally attributed by Veblen to engineers, is neither limited to nor exclusive with that profession. They did, rightfully, insist that the notion of systems theory is not the private preserve of engineering, and they cite Bertanalffy in biology, Churchman in operations research, and R. Buckminster Fuller, wherever he may be. They could have delved into many more streams of intellectual history, as I already have done.[3] They could even have carried their argument a step further and maintained that not only in its origins but in its current applications the usage of the systems approach is not the bailiwick of engineers alone. It has, as I have pointed out elsewhere, become the stock-in-trade of the "brains-for-hire," which often consist of pick-up crews from a motley of professions, not excluding sociologists! If they had made this point, I would have been forced to consider more seriously the phenomenon, "everybody's doing it," which I tried to explain briefly by evoking (and not, as they imply, leaning heavily on) Kuhn's notion of the "dominant paradigm."[4]

Had the authors pursued this course, I might have had to concentrate on what I consider the more basic issues, which do not concern *who* is doing it, nor even *why* they are doing it,[5] so much as what it *is* they are doing, whether it is appropriate, and what are its present and likely effects on planning policy and practices. *What* they are doing cannot be divorced from engineering, for even though the systems approach may have origins as diffuse as its definitions, its technical trappings and its *macromachismo*[6] stem directly from its intimate association, inside or outside of wedlock, with engineering. Appeals to the large wholeness of social phenomena (and hence the need for "systems" thinking) and facile elisions from "systems thinking" to "thinking systematically" notwithstanding, *the systems approach* in practice is a particular, identifiable method, which, I maintain, is neither appropriate nor adequate in the sphere of human and social "systems" understanding, let alone design and management. Rather than provide any more horrible examples, and so as not to leave even the casual reader with any doubts on this matter, I would refer him to Georgescu-Roegen's[7] scholarly refutation of the mechanical analogue in economic process, equally devastating when directed to *all* social process. After showing that logic and rationality, the main articles of faith in the engineers's credo, can handle only a very restricted class of concepts, Georgescu-Roegen asserts that (1) wholesale arithmetization (I prefer the term quantification) is impossible; (2) there is valid knowledge even without arithmetization; and (3) mock arithmetization *is dangerous if peddled as genuine* (my underlining).[8] He then introduces and develops the concept of *entropic indeterminateness*, which is based on thermodynamic principles and which, when applied to systems of economics (and other social activities), throws into sharp perspective the fundamental deficiency inherent in the systems approach as now applied as a nostrum for society's ailments.

It is unfortunate that the authors chose to debase what could have been scholarly discourse by resorting to the *vade mecum* tactic of reinforcing their position by incestuous validation from within the trade, as seen in the quotations from the *IEE* review of my book, *Systems Analysis in Public Policy*.[9] They could, as a matter of fact, have drawn on a more distinguished journal, *Science*,[10] which devoted two pages to a similarly intemperate attack by an engineer connected with "a Program in Engineering for Public Systems." He, in turn, had invoked as validation for his position an angry letter to the editor of *Management Science*[11] by *another* engineer protesting my views on management information systems, which, incidentally, his areospace company employer peddles at a great rate! But this type of exercise can impress only the unenlightened. Such instances serve more to strengthen my contention about the engineers' generally limited world view than to enhance the cause of intellectual advancement. That the book and my widely published writings have received commendation from authorities representing a broad spectrum of disciplines, not excluding engineering, conveys a message more significant than the outraged cries of the wounded species.

---

[2]Lewis Carroll, *Through the Looking Glass*, Grosset & Dunlap, New York, 1946, pp. 230–31.

[3]For more comprehensive treatment of this and other aspects of subject, see Ida R. Hoos, *Systems Analysis in Public Policy: A Critique*, University of California Press, 1972.

[4]I could have stated this another way and cited the prevailing mythology of methodology, persuasive wherein is the message conveyed in the story of the emperor's new clothes.

[5]I was being charitable when I attributed the engineers' quest for social relevance to a possible guilt complex. Despite current trends in government and foundation grantsmanship, I was reluctant to impute this urge to "do the society good" to motives less noble and altruistic, such, perhaps, as opportunism.

[6]Lest I be accused of hyperboie, the following are excerpts from a letter from the University of Illinois Press (October 24, 1973):

Dear Reader:

"How can scientific methods be applied to solving urban problems?
Are problems of traffic congestion, accidents, noise and nuisance solvable?
Should restrictions be placed on privately owned automobiles?
Can close coordination between travel and land use planning be achieved?
Does politics play too great a role in urban planning? Is urban planning really encouraged and sustained for sociological betterment?"

The sales pitch is then made that a "scientific approach to urban analysis," which brings together "scientific techniques from mathematics, operations research, systems analysis, information science, and planning," exists (and in particularly useful form in the books being promoted)!

[7]Nicholas Georgescu-Roegen, *The Entropy Law and the Economic Process*, Harvard University Press, 1971.

[8]*Ibid.*, p. 15.

[9]*op. cit.*

[10]*Science*, Vol. 17, November, 1972, pp. 739–40.

[11]*Management Science*, February, 1972.

Appeals to the claque notwithstanding, the compelling question still remains as to the appropriateness of the tools and techniques being transferred. As I pointed out earlier, Georgescu-Roegen's methodologically refined analysis of the shortcomings and dangers should be understandable in his own terms to the engineer. Moreover, if as the authors seem to imply, there is validity to the *argumentum ad nominem*, one might suggest that they draw on the wisdom of the more talented among their numbers,—none other than Norbert Weiner, who cautioned precisely against this kind of inappropriate transfer, questioned whether it was useful or honest, and even suggested that it was a sham and a waste of time.[12]

Ida R. Hoos
Space Sciences Laboratory
University of California
Berkeley, California
November 1974

Dynamic Systems Analysis in Interdisciplinary Research

ROBERT C. SPEAR.[1] At the 1974 Joint Automatic Control Conference, 39 of the 86 papers presented dealt with topics outside the traditional subject areas of control engineering and dynamic systems theory. In particular, six papers concerned dynamic aspects of biosystems and 28 addressed ecological and environmental subjects. To one who left control engineering to work in the environmental sciences some years ago, the developments signaled by the subject distribution of the JACC papers are of great interest. It would appear that the vigorous development of dynamic systems methodology by engineers and applied mathematicians in the last several decades has now culminated in a widespread recognition of the relevance of the theory to various other disciplines based in the natural sciences. The relevance of dynamic system theory to biology, for example, was historically demonstrated by pioneers like Lotka [1][2] but its broad application and inclusion in teaching programs is a recent phenomenon.

Most of the productive applications of dynamic system methodology can be classified as either scientific or technological in nature. The scientific applications are typified by the writings of Rosen [2] and May [3] which are essentially aimed at the elucidation of the dynamic aspects of scientific problems in biology. The technological or engineering type applications are exemplified in the environmental sciences by the work of Young and Beck [4] in their study of the temporal variations and control of water quality parameters in an English river. A common feature of many of these technological applications of dynamic system theory is that the problem requires the input of various disciplines for its solution. It is often the integrative nature of systems methods that is particularly valuable. However, few problems outside engineering can support an individual who serves only as a dynamic systems analyst. Therefore, to maximize the benefit to be gained from the dynamic systems point of view in interdisciplinary work, it can be argued that we must despecialize and refocus our attention on the problem rather than the methodology used to solve it. There is no doubt that there is increasing scope for dynamic systems analysts in general and control engineers in particular to work in the environmental sciences, the health sciences and in a variety of other interdisciplinary areas. However, for those who choose to pursue this course, new subject matter will have to be learned and the security of a clear technical identity may have to be abandoned. The reward of such a transition is potentially to contribute new insights, different technical viewpoints, and new approaches to interdisciplinary problems.

Underlying this view of the potential role of dynamic systems theory in interdisciplinary work is this writer's experience that its major contribution is strategic rather than tactical. That is, it is the basic perception of the temporal nature of many problems and the ability to conceptualize the processes basic to their understanding in terms of system state, for example, that appear to be the potentially important contributions. One seldom encounters a problem, although they must exist, where knowledge of the best estimation algorithm or a particular theorem from optimization theory is the crucial factor. On the other hand, it is often useful to have at hand some generally useful computer codes for estimation or simulation work.

By way of a concrete example, our group is engaged in a research project aimed at determining the mechanisms by which agricultural field workers can become ill from exposure to foliar residues of certain pesticides while harvesting the crop. Ultimately, such research will provide the basis for the establishment of sound protective strategies. The array of specialized knowledge relevant to the solution of this problem ranges from agricultural chemistry to toxicology with elements of the fields of industrial hygiene, clinical medicine and soil science being of equal importance, to name but a few. However, the key to the problem does not appear to lie in any single specialist's domain. As with many problems in environmental toxicology, it is the integration of the contributions from each discipline into a scheme which displays the interrelationships between variables as well as the dynamics of the system which appear to hold the greatest promise for elucidating the factors important to an understanding of the problem [5]. For example, a determination of the rates at which the pesticide decays and metabolites of toxicological significance are produced on the foliage is of importance as are the effects of environmental variables on these processes. This aspect of the problem lends itself to straightforward differential equation modeling of the form $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)$ $\xi(t), t)$, where $\mathbf{x}(t)$ is the vector of chemical species of toxicological interest in the foliar residue and $\xi(t)$ is a set of parameters associated with environmental variables. The functions $\mathbf{f}$ tend to be poorly defined but in specific cases studied to date the approximations and assumptions usual in engineering practice have been useful. For some of the more important pesticides, it appears that the hazard to workers is a function of the scalar $y = \mathbf{c}^T\mathbf{x}$, where $\mathbf{c}$ is related to the relative toxicity of the chemical species.

Intermediate between these chemical processes and the toxicological processes are a host of variables associated with the exposure mechanism. These variables concern work practices, work rate and other factors similarly difficult to quantify but all of which mediate between the environment and the dose absorbed by the worker. The conceptual or symbolic modeling of these less well-defined processes has been a useful means of integrating the diverse data and, in particular, of delineating the causal relationships between variables.

In common with the environmental phenomena, the absorption of the pesticide through the dermal and respiratory barriers, the kinetics of its binding with the target enzyme system and its eventual excretion from the body are all phenomena for which a temporal interpretation is demanded. However, our inability to quantify the exposure mechanism places an upper bound on the modeling effort that can be justified on the environmental or toxicological components of the problem. There is a natural tendency, nevertheless, to focus on those processes

[12]Norbert Wiener, *God & Golem, Inc.*, 1964, pp. 87–92.

[1]Department of Biomedical and Environmental Health Sciences, University of California, Berkeley, Calif.

[2]Numbers in brackets designate References at end of Forum.

NGR 05-003-329 (23) D7

Journal of Atmospheric and Terrestrial Physics, 1975, Vol. 37, pp. 119–129. Pergamon Press. Printed in Northern Ireland

Properties of the wake of small Langmuir probes on sounding rockets

E. A. BERING
Physics Department, Space Sciences Laboratory, University of California,*
Berkeley, California 94720, U.S.A.

(*Received* 20 *November* 1973; *in revised form* 27 *February* 1974)

Abstract—Split Langmuir probes have been used to study the near wake of small Langmuir probes on sounding rockets. The split Langmuir probe is a device which performs the usual measurement of current versus voltage for the two halves of a probe and, in addition, measures the difference current between the two halves to accuracies the order of 10^{-4} times the single half plate current. Thus it is an ideal instrument for studying the near wake of a small probe. Experiments have been performed on two rocket flights using planar disk and cylindrical geometries and the results presented in this paper. Significant wake perturbations in plasma density and temperature were found due to the probe body itself, even though the probes were the order of or smaller than the Debye length. The largest effects of the wake are seen in the electron collection characteristics of the probe. The wake of small probes show apparent magnetic field aligned structure, even though the probes were much smaller than the ion gyroradius. On one flight, a space charge potential large enough to substantially alter photoemission, $-3 \cdot 5$ volts, was observed.

1. INTRODUCTION

ONE OF the oldest and most fundamental technical problems confronting the experimental space physicist is the interaction of the rapidly moving vehicle and instruments with the ambient medium. Low altitude spacecraft and ionospheric sounding rockets usually have velocities relative to the ambient medium greater than or comparable to the ion thermal velocity. It can be expected that such a spacecraft would have a substantial wake. Numerous experimental results have supported this hypothesis (BORDEAU and DONLEY, 1965; SAMIR and WILLMORE, 1965; HENDERSON and SAMIR, 1967; SAMIR and WRENN, 1969, 1972; WRENN, 1969; SAMIR, 1970; OYA, 1970; MILLER, 1972; SAMIR *et al.*, 1973). A review of the problem was given by LIU (1969).

Several approximate theoretical models have been made of the problem, but a quantitatively adequate general solution has not yet been attempted due to the difficulty of the problem (GUREVICH *et al.*, 1969; SAMIR and WRENN, 1969; BOGASHCHENKO *et al.*, 1971; SAMIR and JEW, 1972). Experimenters attempting to make probe measurements of ionospheric plasma parameters have had to use extreme care in treating the wake problem when interpreting and analyzing their data. One approach is to assume that a boom mounted probe, comparable to or smaller than the Debye length in at least one dimension, many Debye lengths from the spacecraft, and oriented at right angles to the direction of travel, is not affected by wake problems (SAMIR and WILLMORE, 1965; FINDLAY and BRACE, 1969; BRACE *et al.*, 1971). The wake of the probe is typically ignored. The purpose of this paper is to present data on the structure of the wake of probes with dimensions on the order of the Debye length flown on two sounding rockets.

* Now at: Department of Physics, University of Houston, Houston, Texas, 77004, U.S.A.

The Instrument

The instrument used for this study was a split Langmuir probe (Bering *et al.*, 1973a,b), which consists of two parallel conducting plates or two opposite conducting hemicylinders, separated by an insulator. The plates are maintained at the same potential, which is periodically swept. The current to one plate, I_1, and the difference current between the plates, $I_1 - I_2$, are measured as functions of the plate voltage. Two such probes were flown on each rocket, one mounted with the normal to the probe surface perpendicular to the spin axis, and the other, mounted with its normal parallel to the spin axis. The probes were deployed on ½ meter booms mounted at the forward end of the rocket payload. The results from the first of the two flights discussed here and the details of how the instrument works have been discussed in two previous papers (Bering *et al.*, 1973a,b) called papers I and II respectively.

On the two flights discussed in this paper, the voltage sweep rate of the perpendicular probe was fast compared to the spin rate. Thus, the plasma density and temperature deduced from the current characteristic of one plate can be analyzed as a function of spin phase. The directional nature of the single plate of the split probe gives the instrument the capability of detecting its own wake as the rocket spins. It is assumed that forward mounted, boom deployed probes are not affected by the vehicle wake, at least on the upleg. Therefore, any wake phenomena detected can be attributed to the probe itself.

Analysis of the difference current as a function of potential and spin phase can yield further information about wake structure. As shown in paper II, the difference current at positive or negative potentials is given by

$$\Delta I = neA\hat{a} \cdot \bar{v}_i \qquad \text{for } V < 0 \tag{1}$$

$$\Delta I = neA\hat{a} \cdot \bar{v}_e \qquad \text{for } V > 0 \tag{2}$$

where ΔI is the difference current, n is the free stream plasma density, A is the probe area, e is the elementary charge, $\hat{a}$ is a unit vector normal to the probe, $\bar{v}_e$ and $\bar{v}_i$ are the electron and ion flow velocities respectively, and V is the probe potential. If $\bar{v}_e$ and $\bar{v}_i$ are known from other data, and analysis of ΔI does not give agreement, the discrepancies are interpreted as being due to wake associated phenomena, or wake effects. The first order results given in equations (1) and (2) are also effects of the flow of plasma past the probe, but are not usually described as wake effects.

3. Flight 1—Black Brant VB-18

The first flight of the split Langmuir probe was on a Black Brant VB sounding rocket launched from Fort Churchill on 2 August 1968 at a time near local midnight. The results of the flight and the details of the instrument have been reported in papers I and II.

The probes on this flight were aquadag coated, flat, circular discs, 8 cm in diameter and 1 mm in thickness and which were mounted on booms approximately ½ meter from the rocket body. By comparison at apogee the Debye length was 1·2 cm, the thermal electron gyroradius was 2·2 cm, and the thermal ion gyroradius

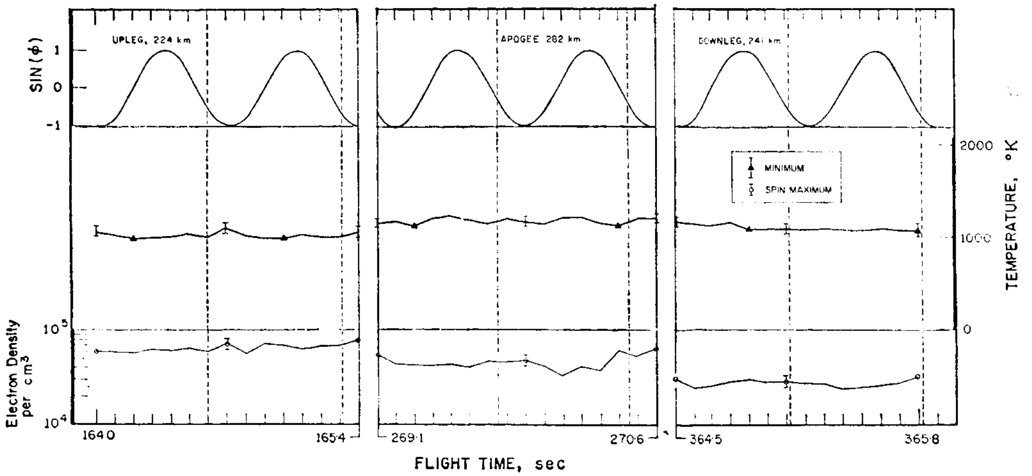


Fig. 1. Three high time resolution samples of the temperature and density measurements deduced from data from the perpendicular detector on the Black Brant rocket.

4 meters. Raw flight data from the perpendicular detector was discussed in papers I and II. The superficial discussion of wake phenomena in the two previous papers concluded that the raw data indicated that the apparent electron flow velocity, as indicated by the difference current at positive potentials, was much larger than the apparent ion flow, as indicated by the difference current at negative potentials. This anomaly was interpreted as being due to probe wake. It is discussed further below.

As discussed above, the single plate current data can be analyzed to give information of plasma temperature and density in the wake of a probe near plasma potential. Three examples are shown in Fig. 1, which shows the sine of the spin phase angle, the electron temperature, and the electron density as measured by the perpendicular probe. The temperature data shown in Fig. 1 is typical of the entire flight in that there is no significant systematic spin variation of the measured temperature. The dashed lines indicate the location of the density maxima expected from the times of maxima of a sine wave fitted to the ion difference current data. Within the errors, there is good agreement between the times of the density maxima from the difference current data, and the density maxima in the single plate current data. It can be seen from the figure that the lowest apparent density observed on a given spin was typically 80–85 per cent of the maximum observed that spin. This is interpreted as a decrease in plasma density in the wake of the probe itself. This perturbation cannot be due to the vehicle wake because, in such a case, the single plate maxima would be 90° out of phase with the difference current maxima. In a situation where vehicle wake is the dominant wake influence in the probe, the single plate maxima occur when the probe is farthest from the wake, i.e. when the boom is pointing foreward along the velocity vector. On the other hand, the difference current maxima still occur when the normal to the plate is most parallel to the velocity. It is significant that little apparent change in the perturbation

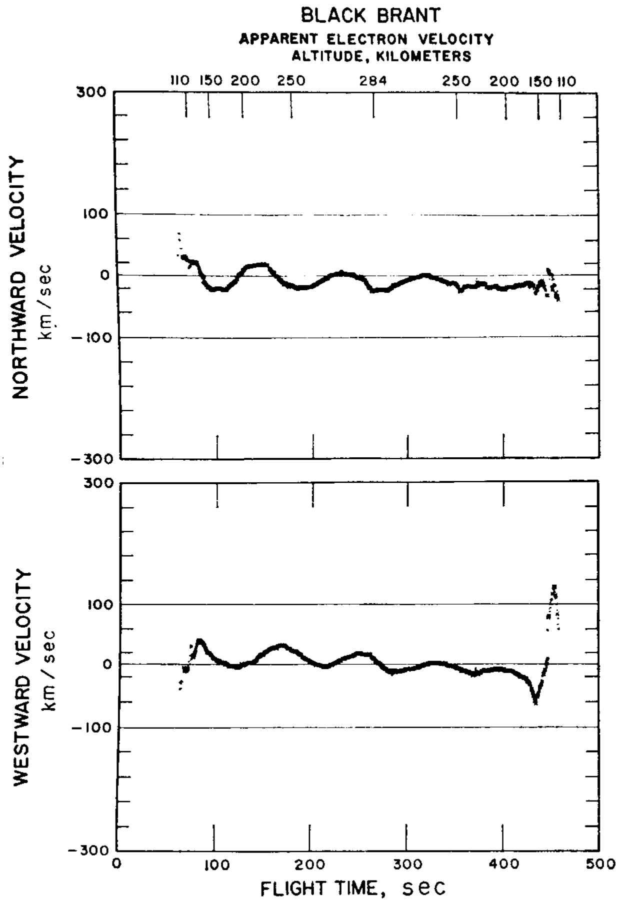


Fig. 2. Apparent electron flow velocity plotted in Earth fixed geomagnetic coordinates. These data are believed to be manifestations of the wake, not the actual electron flow velocity.

was observed on the downleg. This implies that the probe was observing only its own wake throughout much of the flight.

The methods of analyzing the difference current data have been discussed in paper I. Figure 2 shows the apparent electron flow velocity in geomagnetic coordinates. As discussed in paper II, the large magnitude of the apparent electron velocity was an effect of the wake of the probe.

This data has a number of features which reveal aspects of the structure of the probe wake. The magnitude of the apparent flow velocity varied periodically from near zero to a maximum that was typically 10–15 per cent of the electron thermal velocity. The period of the variation was the same as the precession period of the vehicle, and the magnitude of the apparent flow velocity was a minimum at the same time that the angle between the rocket spin axis and the magnetic field was its minimum value of 7°. The interpretation of Fig. 2 is that the large difference current is due to a density decrease in the wake of the probe. Consequently, the wake facing side of the probe receives a substantially reduced thermal flux of plasma. Since the electron thermal velocity is much larger than the plasma bulk flow velocity, the wake effect difference current is much larger than the first order linear difference current as given in equation (2).

Furthermore, it is apparent that the density decrease is a function of rocket attitude, and increases as the angle between the rocket spin axis and the magnetic field increases. This could be due either to a field aligned wake structure that appears in the difference current as the rocket tips over, to a wake increase caused by an increase in the projected area of the probe normal to B, or to a combination of the two.

It is also significant that the magnitude of the density decrease at positive potential is approximately the same as the decrease observed at plasma potential. This indicates that the effect of the probe potential on the structure of the wake is relatively small for this geometry.

An analysis of the difference current data at negative potentials from this flight has been reported in paper I. In that paper, a comparison between the electric field deduced from the measured ion flux, and the electric field reported by KELLEY *et al.* (1971) from a double probe experiment on the same rocket was made. The initial comparison showed significant precession periodic discrepancies between the two measurements. In order to reduce these discrepancies, a first order correction for wake effects was made, as described in paper II. Figure 3 shows a comparison of the electric field deduced from uncorrected and corrected data, and the electric field reported by KELLEY *et al.* (1971). The figure shows that a small reduction in the discrepancies has thereby been achieved. This indicates that there are basic similarities in the structure of the probe wake at positive and negative potentials. The magnitude of the remaining discrepancies argues that there are some equally fundamental dissimilarities as well.

Furthermore, the differences between the two measurements are larger on the downleg than on the upleg and are largest near apogee. Since the vehicle wake is most likely to influence a foreward boom mounted probe at apogee, and more likely to influence such a probe on downleg than up, this may indicate that the major disturbance is due to the vehicle wake, rather than the wake of the probe itself. Such an effect would not be reduced by the correction mentioned above, which corrects only for the self wake of the probe. This interpretation is supported by the observation that the raw difference current data at negative probe potential deviates most from sinusoidal at times when the disturbance is maximum.

The effect of this deviation is further discussed in paper II, where a comparison of the electric field results deduced by analyzing data from two different negative probe potentials is presented. Paper II shows that there are significant differences between the two results. particularly around 260 and 400 sec flight time. They were attributed either to the effect of the vehicle wake on probe function, or to variations in the field aligned structure in the probe's wake as a function of probe potential.

There were also indications of a substantial space charge in the wake of the probe. It was shown in papers I and II that photoemission from the probes during

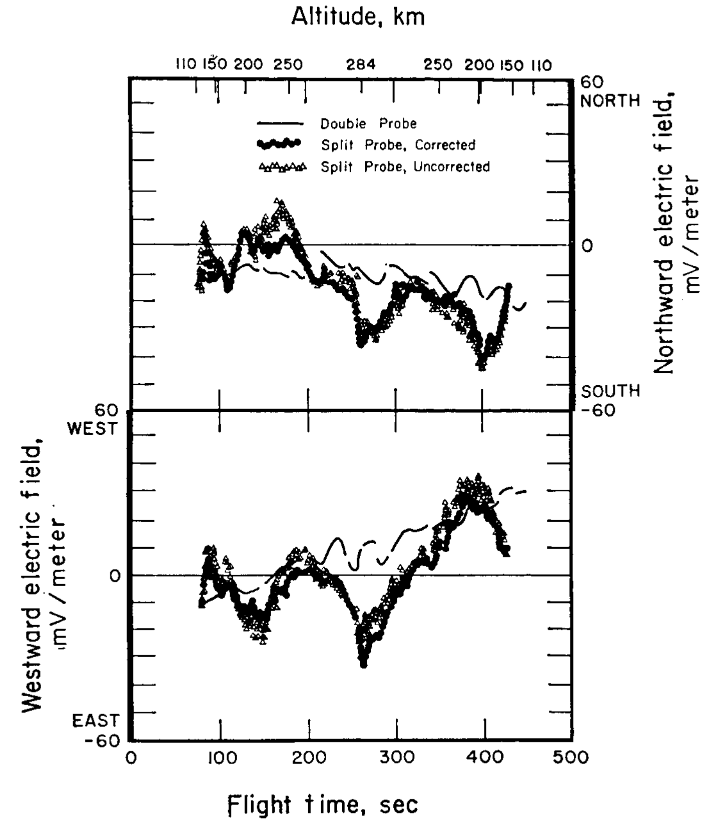


Fig. 3. Perpendicular electric field components in the Earth fixed geomagnetic north and west directions deduced from data taken by the Black Brant payload. The figure shows the results of applying a wake effect correction to the split Langmuir probe data.

the flight was reduced by a factor of 30, presumably by wake effects.

This observation was checked by including the expected photoemissivity in the ion flow analysis. The result is shown in Fig. 4. The gross disagreement between the deduced electric field shown in the figure and the electric field reported by KELLEY *et al.* (1971) and shown above is confirmation that the photoemission was in fact substantially reduced.

Examination of the photoelectron spectrum reported by HINTEREGGER (1960) indicates that photoemission reduction of the observed magnitude would result from a space charge potential of $-3{\cdot}5$ volts in the wake.

4. JAVELIN 8·56

The second successful flight of the split Langmuir probe was on a Javelin sounding rocket launched on 3 April 1970 from Fort Churchill, Manitoba, Canada, at a time near local midnight. The probes on this flight were 3 mm diameter cylinders. The split section was 9·5 mm long, with a 14 mm long guard cylinder at each end driven at the same potential as the detecting hemicylinders. The probes were mounted on one meter booms at the nose of the vehicle.

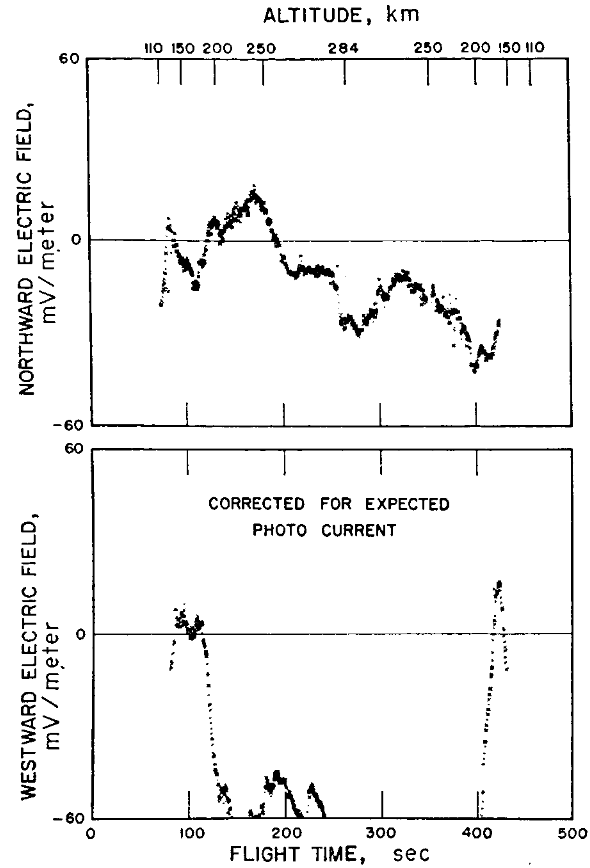


Fig. 4. Perpendicular electric field components in the Earth fixed geomagnetic north and west directions deduced from data corrected for photoemission of nominal intensity. Data taken by the Black Brant payload.

On this flight, there was a malfunction somewhere in the payload which had the effect of driving the payload potential negative with respect to the plasma. As a result, the split Langmuir probe input electronics went out of common mode range about 400 sec into the flight. During the period of the flight that good data was obtained, the Debye length varied from 2 mm to 1 cm, the thermal electron gyroradius was about 2 cm and the thermal ion gyroradius about 4 meters.

The rocket attitude on this flight was exceptionally stable. The precession cone half angle was only 2·8°. For the purpose of analysis, it has been assumed that a precession cone this small introduces negligible precession dependence in the wake. Under this assumption, each density value measured by the perpendicular probe at plasma potential has been normalized with respect to the density corresponding to the maximum flux measured at plasma potential on any sweep during that particular spin. The normalized values corresponding to a given number of sweeps away from the maximum were then averaged over the entire body of good data. A similar analysis has been performed for electron temperature, normalized

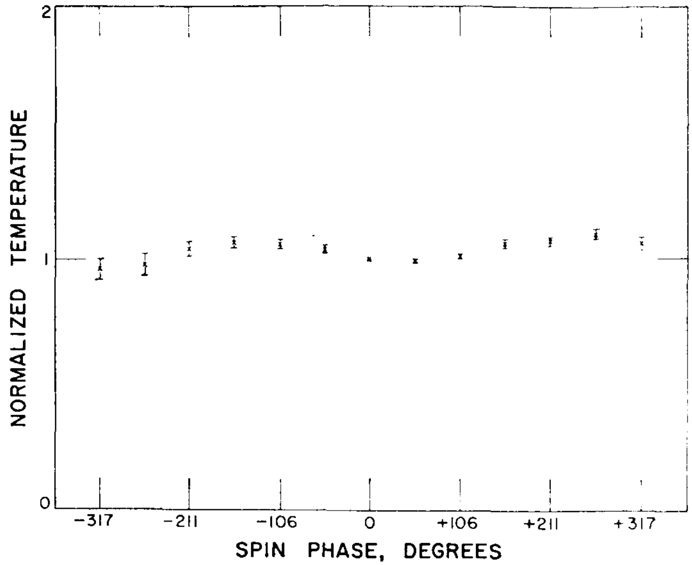


Fig. 5. Normalized, averaged electron temperature measured by the Javelin payload and plotted as a function of spin phase.

with respect to the temperature measured on the same sweep as the density maximum that spin. The results of this analysis are shown in Figs. 5 and 6. The error bars shown are the errors in the mean.

The curve shown in Fig. 5 indicates a slight but significant rise in the temperature in the probe wake, on the average. This is consistent with results reported by

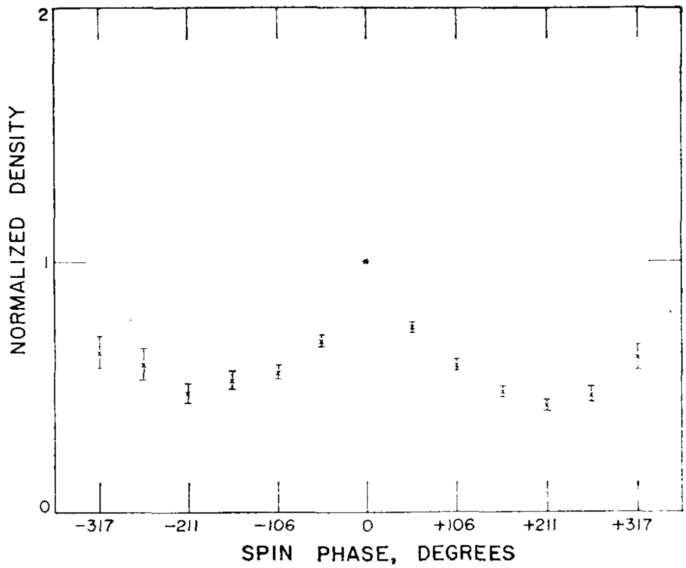


Fig. 6. Normalized averaged electron density measured by the Javelin payload and plotted as a function of spin phase.

SAMIR and WRENN (1972). It is interesting that this temperature increase can be observed in the wake of an object as small as these probes.

The curve shown in Fig. 6 indicates an approximately 50 per cent decrease in the wake of the probe. This was not expected for a probe this small. One of the design hypotheses behind the idea of flying a small probe was the idea that the wake of a small probe at plasma potential would be small. A 50 per cent decrease in the density of a small probe was unexpectedly large. The result raises a substantial question about the absolute accuracy of all spacecraft Langmuir probe analyses done under the assumption that small probes mounted well clear of the body wake are not significantly affected by their own wake.

The common mode problem mentioned above made accurate determination of the apparent electron velocity impossible, since the common mode problem had the effect of truncating and distorting the positive portion of the sweep. As a result, accurate determination of the amplitude and phase of the sinusoidal variation of the difference current at positive potential was impossible. Hand analysis of randomly selected points throughout the flight gave apparent electron flow velocities of between 40 and 80 km/sec. This is about 20–40 per cent of the electron thermal velocity, and indicates a wake density decrease of the same amount. The reason for the apparent reduction in the wake density decrease at positive potential as compared to plasma potential is not understood. This apparent reduction in the decrease could be a spurious effect introduced by the common mode error instead of being a real property of the wake of the probe. In either case, 20–40 per cent can be taken as a lower limit on the decrease in plasma density in the rear wake of a small cylindrical probe at positive potential.

In the absence of good data at positive potential, the technique employed in the previous section for discussing the properties of the wake at negative potential cannot be used. This difficulty is compounded by the fact that the payload's electronic maladies had the additional unfortunate effect of introducing substantial commutator noise on the high gain difference current channel of the perpendicular split probe. As a result, the data had to be averaged over 16 spin periods to produce meaningful results. The averaged split probe data and data from a double probe experiment on the same rocket (KELLEY, private communication), are shown in Fig. 7. All that can be deduced from this figure about the wake at negative potential is that the density decrease must be much less than 50 per cent probably on the order of 10–20 per cent at most.

5. CONCLUSIONS

A new technique for examining the structure of the wake of small Langmuir probes on spacecraft has been discussed. The main conclusions that can be drawn from this study are:

(1) Small probes, even probes smaller than the Debye length, do have substantial wakes.

(2) The largest effects of the wake are seen in the electron collection characteristics of the probe. Since these rockets were moving slowly compared to the electron thermal velocity, the wake is presumed to be primarily due to the interaction of the ions with the vehicle. Therefore, the electron wake is due to the effect of local

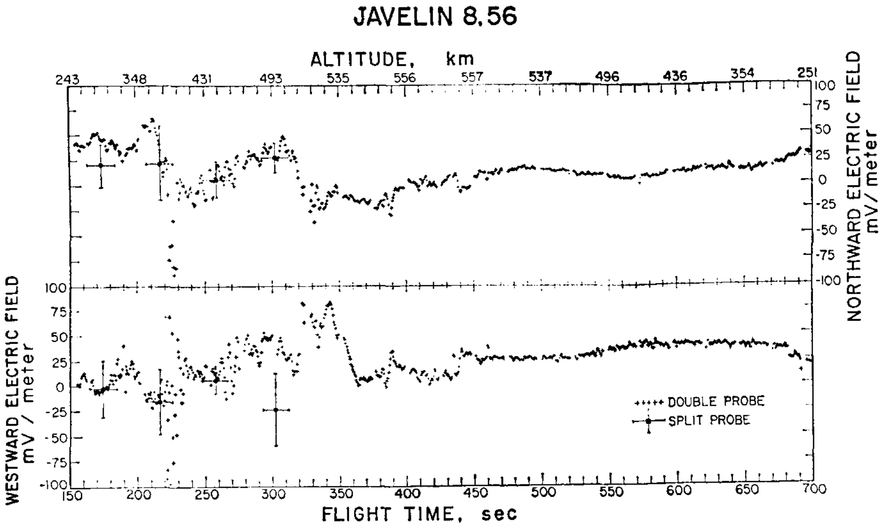


Fig. 7. Perpendicular electric field components in the Earth fixed geomagnetic north and west directions measured by two experiments on the Javelin payload.

electric fields. This implies that any accurate wake model must include electric field terms.

(3) The wakes of small probes show apparent magnetic field aligned structure. This is an unexpected result, considering the large ratio of ion gyroradius to probe radius. This result places further restraints on allowed model approximations.

(4) The observation that small probes do have substantial self wakes calls into question one of the key assumptions used in analyzing spacecraft Langmuir probe data. This is the assumption that small, boom mounted probes do not suffer from selfwake problems.

(5) Space charge potentials in wakes can be large enough to substantially alter photoemission.

Acknowledgements—The author wishes to thank Professor F. S. Mozer for his many helpful comments and ideas, and for reviewing the manuscript. He also wishes to thank Dr. M. C. Kelley for helpful comments and the double probe electric field data, and Dr. U. V. Fahleson for helpful comments. Finally, he also wishes to thank Dr. A. Kavadas, Mr. M. Hodson, Mr. N. Foulds and the Space Engineering Division of the University of Saskatchewan, Canada, for the assistance they provided during the Black Brant program.

The instrument development and data analysis required for these programs were supported by the Atmospheric Sciences Section of the National Science Foundation under grants GA-918 and GA-1317, by the National Aeronautics and Space Administration under grants NGR 05-003-239 and NGR 05-003-445 and in part by a contribution from the National Research Council of Sweden. Space was provided on the Black Brant by the National Research Council of Canada.

References

Bering E., Kelley M. C. and Mozer F. S. 1973a *J. geophys. Res.* **78**, 2201.

Bering E., Kelley M. C., Mozer F. S. and Fahleson U. V. 1973b *Planet. Space Sci.* **21**, 1983.

Bogashchenko I. A., Gurevich A. V., Salimov R. A. and Eidelman Yu. I.	1971	*Sov. Phys. JETP* **32**, 841.
Bourdeau R. E. and Donley J. L.	1965	*Proc. R. Soc.* **A281**, 285.
Brace L. H., Carignan C. R. and Findlay J. H.	1971	*Space Research XI*, p. 1079. Akademie Verlag, Berlin.
Findlay J. A. and Brace L. H.	1969	*Proc. IEEE* **57**, 1054.
Gurevich A. V., Pitaevskii L. P. and Smirnova V. V.	1969	*Space Sci. Rev.* **9**, 805.
Henderson C. L. and Samir U.	1967	*Planet. Space Sci.* **15**, 1459.
Hinteregger H. E.	1960	*Space Research I*, p. 304. North-Holland, Amsterdam.
Kelley M. C., Mozer F. S. and Fahleson U. V.	1971	*J. geophys. Res.* **76**, 6054.
Liu V. C.	1969	*Space Sci. Rev.* **9**, 423.
Miller N. J.	1972	*J. geophys. Res.* **77**, 2851.
Oya H.	1970	*Planet. Space Sci.* **18**, 793.
Samir U.	1970	*J. geophys. Res.* **75**, 845.
Samir U. and Willmore A. P.	1965	*Planet. Space Sci.* **13**, 285.
Samir U. and Jew H.	1972	*J. geophys. Res.* **77**, 6819.
Samir U., Maier E. J. and Troy B. E., Jr.	1973	*J. atmos. terr. Phys.* **35**, 513.
Samir U. and Wrenn G. L.	1969	*Planet. Space Sci.* **17**, 693.
Samir U. and Wrenn G. L.	1972	*Planet. Space Sci.* **20**, 899.
Wrenn G. L.	1969	*Proc. IEEE* **57**, 1072.

D8

(27)

Finite Heat Conduction Effects On Ion Cyclotron and Drift Cyclotron Modes

Mary K. Hudson

Space Sciences Laboratory
University of California
Berkeley, California 94720

January 1975

Abstract

Finite heat conduction effects on the ion cyclotron and drift cyclotron instabilities are investigated, along with the converse effect of these modes on electron thermal conductivity.

Introduction

Recent Q machine experiments have demonstrated that the high frequency electrostatic ion cyclotron[1] (IC) and ion cyclotron drift[2] (ICD) instabilities generate anomalous resistivity which inhibits the destabilizing current when the plasma is stable to the ion acoustic mode; apparently the low frequency drift instability[3] does not[4]. Finite heat conduction affects on the ion acoustic[5] and drift[6,7,8] modes have been investigated and shown to be destabilizing in certain regions of parameter space. It is the purpose of this paper to analyze the effect of finite electron thermal conductivity on the IC and ICD modes. Electron thermal conductivity along a z-directed magnetic field is reduced by collisions at parallel wavelengths exceeding the electron mean free path $k_z\lambda_e<1$. In a turbulent plasma the wave-particle interactions which produce anomalous resistivity can either increase or decrease the thermal conductivity, depending on the mode which is generating the turbulence.[9] After treating the effects of finite electron heat conduction on the linear IC and ICD modes, we will consider the influence of these modes on anomalous thermal conductivity.

Linear Dispersion Relation

Hendel and Yamada[2] have derived the linear dispersion relation for the IC and ICD modes in a fully ionized, collisionless, inhomogeneous plasma from the Vlasov equation. In these modes with $k_z/k_\perp <<1$, the electrons move primarily parallel and the ions perpendicular to the magnetic field. Collisions should be included in the electron contribution χ_e to the dielectric function

$\varepsilon = 1 + \chi_i + \chi_e$ for $k_z\lambda_e<1$, realized in some experiments.[4] The ions remain collisionless for $k_\perp\rho_i\geq1$ and $\rho_i/\lambda_e<<1$, where ρ_i is the ion Larmor radius. To include collisions, and in particular finite electron thermal conductivity, we compute χ_e from the electron fluid equations of continuity, momentum and heat transfer in the ion frame, which are valid for $k_z\lambda_e<1$[10]

$$\partial n/\partial t + \nabla \cdot (n\underline{v}) = 0 \tag{1}$$

$$mn(d\underline{v}/dt) = -\nabla(nT_e) - ne[\underline{E} + (\underline{v}/c) \times \underline{B}] - C_t n\nabla T_e + \underline{R}_{ei} \tag{2}$$

$$3/2\ n(dT_e/dt) + nT_e\nabla\cdot\underline{v} = -\nabla\cdot q + Q \tag{3}$$

$$R_{ei} = -C_r mn\nu_e \underline{v} \tag{4a}$$

The coefficients C_r=0.51, C_t=0.71, and C_x=3.16. The parallel electron heat flux is

$$\underline{q} = C_t nT_e v_z - C_x(nT_e/m\nu_e)\nabla T_e \tag{4b}$$

The collisional energy transfer from electrons to ions is given by

$$Q = -3(m/M)n\nu_e(T_e - T_i) \tag{4c}$$

Both depend on the electron-ion collision frequency ν_e.

We neglect electron inertia at low frequencies ($\omega<<\omega_{pe}$), and treat the perpendicular electron motion in the guiding center approximation for $k_\perp/k_z>>1$.[8] Retaining the collisionless Vlasov result for χ_i, the dispersion relation becomes

$$\varepsilon = 1 + \chi_e + \chi_i = 0 \tag{5a}$$

$$\chi_i = \frac{R_T}{k^2\lambda_D^2}\left\{1 - \Sigma \frac{\omega - R_T\omega_D}{\omega - n\Omega}\Gamma_n \left[1 - i\sqrt{\pi}\frac{\omega-n\Omega}{k_z a_i}\exp\left(-\frac{(\omega-n\Omega)^2}{k_z^2 a_i^2}\right)\right]\right\}$$

$$\chi_e = \frac{1}{k^2\lambda_D^2}\,\frac{\omega_D(\omega-C_t''\omega_1) + i\nu_\parallel(\omega+\bar{\chi}\omega_D+C_t'\omega_1) - \nu_\parallel^2(\bar{\chi}-\bar{\xi})}{\omega(\omega-C_t''\omega_1) + i\nu_\parallel\left[\omega(1+\bar{\chi}) + 1/3\, C_t'\omega_1\right] - \nu_\parallel^2(\bar{\chi}-\bar{\xi}) + \omega_1^2}$$

Here $\omega_D/k_y = (cT_e/eB)\nabla(\ln n)$ is the electron diamagnetic drift in an x-directed density gradient, $\omega_1/k_z = V$ is the current-driven electron drift, and $\nu_\parallel = k_z^2 a_e^2/\nu_e$ is the parallel electron heat conduction rate. $\Gamma_n = I_n(b)e^{-b}$ where $b = k_\perp^2 a_i^2/2\Omega^2$; $C_t' = (1 + C_t)$, $C_t'' = (2 + C_t)$, $\bar{\chi} = 2/3\left[C_r C_x + (1+C_t)^2\right]$, $\bar{\xi} = 2/3\,(1 + C_t)^2$ and $R_T = T_i/T_e$ is the ion-electron temperature ratio. Other standard notation includes the electron and ion thermal speeds a_e and a_i, ion cyclotron frequency Ω and electron Debye length λ_D. We have expanded the ion Z-function in (5a) for large argument $(\omega - n\Omega)/k_z a_i \gg 1$ and henceforth assume $k\lambda_D \ll 1$ for the modes of interest.

We will compare the fluid electron dielectric in (5a) with the collisionless Vlasov result[2]

$$\chi_e = \frac{1}{k^2\lambda_D^2}\left[1 + i\sqrt{\pi}\,\frac{\omega-k_zV-\omega_D}{k_z a_e}\right] \tag{5b}$$

Here the electron Z-function has been expanded in the small argument limit for $(\omega - k_zV)/k_z a_e \ll 1$.

Marginal Stability

At marginal stability we can set the real and imaginary parts of (5a) separately equal to zero and solve the real part for the frequency at the $n = 1$ harmonic

$$\omega = \Omega \left\{ 1 + \frac{\Gamma_1(b)(1-R_T\omega_D/\Omega)}{1-G(b)+R_T \dfrac{\omega_D\Omega-\nu_\parallel^2(\overline{\chi}-\overline{\xi})]}{[\Omega^2-\nu_\parallel^2(\overline{\chi}-\overline{\xi})]}} \right\} \equiv \Omega(1+\Delta) \tag{6a}$$

where $G(b) \equiv \Gamma_1 + (1-\Gamma_o)/b$.[11] The isothermal result[2] is recovered as $\nu_\parallel(\overline{\chi}-\overline{\xi}) \to \infty$ $(\nu_e \to o)$.[8] For $\omega_D < \Omega (> \Omega)$ the parameter Δ is increased (decreased) by the finite heat conduction terms. This corresponds to an increase (decrease) in the marginally stable electron drift obtained from the imaginary part of (5a), assuming $\nu_\parallel >> \omega$, ω_1 and neglecting ion Landau damping for $(\omega-n\Omega)/k_z a_i >> 1$[16]

$$\frac{V}{a_e} > \left(\frac{\omega-\omega_D}{k_z a_e}\right)\left(\frac{\overline{\chi}}{(\overline{\chi}-\overline{\xi})}\ \frac{1}{k_z\lambda_e}\right) \Bigg/ \left[\frac{2/3}{\sqrt{\pi}}\ \frac{(1+C_t)}{(\overline{\chi}-\overline{\xi})}\ \frac{1}{k_z\lambda_e}\right] \tag{6b}$$

In contrast to (6b), the collisionless kinetic threshold electron drift obtained by substituting (5b) for χ_e in (5a) is[2]

$$\frac{V}{a_e} > \left(\frac{\omega-\omega_D}{k_z a_e}\right) \tag{6c}$$

The combined threshold is

$$\frac{V}{a_e} > \left(\frac{\omega-\omega_D}{k_z a_e}\right)\left(1 + \frac{1}{\sqrt{\pi}}\ \frac{\overline{\chi}}{(\overline{\chi}-\overline{\xi})}\ \frac{1}{k_z\lambda_e}\right) \Bigg/ \left[1 + \frac{2/3}{\sqrt{\pi}}\ \frac{(1+C_t)}{(\overline{\chi}-\overline{\xi})}\ \frac{1}{k_z\lambda_e}\right] \tag{6d}$$

We have included both fluid and resonant destabilization in (6d). The two are derived separately for fluid and Vlasov electrons, respectively, and only one applies in a given limit of $k_z\lambda_e$. However, this is accounted for in (6d) since the fluid result dominates for $k_z\lambda_e << 1$. In the limit $k_z\lambda_e << 1$, (6d) demonstrates that finite electron thermal conductivity $(\overline{\chi})$ is required to couple the destabilizing current to the IC mode, just as it is for the collisional, current driven ion acoustic[5] and drift[7] modes. Thus, while the threshold drift is increased by about a factor of two or over the collisionless limit, it is finite heat conduction that enables the IC mode to persist in the collisional case.

For $\omega_D > \omega$, the current threshold in (6d) vanishes, and (6a) yields the nonresonant ICD mode destabilized directly by the density gradient. The instability threshold, $\omega_D > \omega$, is reduced as Δ is reduced by finite heat conduction for $\omega_D > \Omega$ in (6a). While the IC mode is a current driven instability with drift corrections in an inhomogeneous plasma, the ICD mode is destabilized directly by the density gradient for scale lengths the order of the ion Larmor radius. This latter mode persists in the collisional limit, even for infinite thermal conductivity ($\chi \to \infty$ in 6d). Likewise, the collisional drift[6,7,8] and ion acoustic[5] modes have non-current driven branches which persist in the isothermal approximation.

A further comparison can be made with finite heat conduction effects on the collisional drift mode destabilized by the ion finite Larmor radius drift in the absence of a parallel current[6,7,8]. Hudson and Kennel[8] have shown that finite heat conduction increases (decreases) the drift instability threshold when $k_\perp/k_z$ is small (large)[14]. Likewise, the fastest growing IC mode has $k_\perp/k_z \sim 10$, and the threshold is increased by finite heat conduction. However, the fastest growing ICD mode has $k_\perp/k_z >> 10$, and the threshold is decreased. Finite parallel electron thermal conductivity appears to enhance strongly field alligned modes ($k_\perp/k_z >> 1$).

Nonlinear Effects

Electrical and thermal conductivity are finite to the same order of approximation in the fluid equations (1-3) which are moments of the Vlasov equation below. In a turbulent plasma the wave-particle interactions which produce anomalous resistivity can either increase or decrease the thermal conductivity, depending on the mode which is generating the turbulence[9]. Caponi and Krall[9] investigated the effect on parallel electron thermal conductivity of modes with $k_z/k_\perp \ll 1$ destabilized by a perpendicular current. Their procedure was to take moments of the Vlasov equation, assuming a quasilinear form for the nonlinear term on the right hand side[12],

$$\left[\frac{\partial}{\partial t} + \underline{v} \cdot \frac{\partial}{\partial \underline{x}} + \frac{F}{m} \cdot \frac{\partial}{\partial \underline{v}}\right] f(\underline{x},\underline{v},t) = \frac{-q}{m}\frac{\partial}{\partial \underline{v}} \left\langle \tilde{E}\, \tilde{f} \right\rangle \tag{7}$$

where $\tilde{E}$ and $\tilde{f}$ are the perturbed electric field and particle distribution function.

While the $\underline{k}$ spectrum is 3-dimensional, only spatial variation and heat flow in the z-direction are considered. The electron distribution function is assumed to be quasimaxwellian, $f = f_0\,(z, (\underline{v} - \underline{V})^2, t) + f_1\,(z, v_z, (\underline{v}-\underline{V})^2; t)$, yielding a complete set of three equations in v_z, nT_z and the random heat flux $Q_z = m/2\int dv\, v_z^{\,3} f_1$. The set is reduced to two equations by assuming ambipolarity, $\partial/\partial t(v_z^{\,e} - v_z^{\,i}) = 0$. The equation governing the anomalous thermal conductivity on the right hand side

$$\frac{\partial Q_z}{\partial t} + \frac{3}{2}\, n\, \frac{T_z}{m}\frac{\partial T_z}{\partial Z} = \frac{3}{2}\, q \left\langle \tilde{E}_z \left(\tilde{v}_z^{\,2} n - \frac{T_z}{m}\tilde{n}\right)\right\rangle = 3q\int d\underline{k}\, \mathcal{E}_{\underline{k}}\, k_z \int d\underline{v} \left(v_z^{\,2} - \frac{T}{m}\right) \mathrm{Im}\,\epsilon_{\underline{k},\underline{v}} \tag{8}$$

is solved simultaneously with a heat transfer equation like (3), which now includes a turbulent heating term. Here the electric field energy density obeys $E_{\underline{k}}(t) = E_{\underline{k}o}\, e^{2\gamma_{\underline{k}} t}$, where $E_{\underline{k}o}$ is the thermal noise level[13] and the growth rate $\gamma_{\underline{k}}$ is determined by the linear dielectric function $\varepsilon_{\underline{k}} = \int d\underline{v}\ \varepsilon_{\underline{k},\underline{v}}$.

Of the three modes which they investigated (Buneman, modified two-stream and ion acoustic) only the nonresonant Buneman satisfied $v_z^{\,2} > T_e/m$ or $\omega/k_z a_e > 1$ in (8), leading to a negative anomalous thermal conductivity or enhanced heat loss. Physically, the Buneman mode propages along the magnetic field faster than electrons, and conducts energy away faster. We, therefore, expect other nonresonant modes like the ICD, which has $\omega/k_z a_e \gg 1$, to increase electron thermal conductivity. On the other hand, finite heat conduction effects on the resonant IC mode, which has $\omega/k_z a_e \ll 1$, should be more pronounced. Evaluation of the anomalous thermal conductivity matrix element for the case of a parallel current, which destabilizes the IC mode, and a perpendicular density gradient, which destabilizes the ICD mode, will be presented elsewhere.

Summary

In conclusion, we have investigated the effects of finite parallel electron thermal conductivity on the current (IC) and density gradient (ICD) driven ion cyclotron modes. We have found that finite heat conduction increases the IC threshold and decreases the ICD threshold. The nonlinear effect of these modes on thermal conductivity tends toward stability, as one would expect. From physical arguments it follows that the nonlinear IC mode decreases while the ICD mode increases parallel electron heat loss. The latter effect may limit ion heating[1,15] if a local nonlinear mechanism exists which transfers perpendicular ion to parallel electron energy.

References

[1]Benford, G., N. Rynn, J.J. Thompson and W.S. Williamson, Phys. Fluids 17, 101, 1974.

[2]Hendel, H.W. and M. Yamada, Phys. Rev. Lett., 33, 1076, 1974.

[3]Coppi, B. and E. Mazzucato, Phys. Fluids, 14, 134, 1971.

[4]Hendel, H.W. and M. Yamada, private communication, 1974.

[5]Rognlien, T.D. and S.A. Self, Phys. Rev. Lett., 27, 792, 1971.

[6]Tsai, T.S., F.W. Perkins and T.H. Stix, Phys. Fluids, 13, 2108, 1970.

[7]Ellis, R.F. and R.W. Motley, Phys. Fluids, 17, 582, 1974.

[8]Hudson, M.K. and C.F. Kennel, submitted to J. Plasma Phys., 1974.

[9]Caponi, M.Z. and N.A. Krall, Bull. Am. Phys. Soc., 19, 925, 1974.

[10]Braginskii, S.I., Reviews of Plasma Physics, I, M.A. Leontovitch ed., Consultants Bureau, New York, 205, 1965.

[11]Kindel, J.M. and C.F. Kennel, J. Geophys. Res., 68, 2579, 1963.

[12]Liewer, P.C. and N.A. Krall, Phys. Fluids, 16, 1953, 1974.

[13]Krall, N.A. and A.W. Trivelpiece, Principles of Plasma Physics, McFraw-Hill New York, Ch. 8 and 11, 1973.

[14]The drift instability threshold is increased (decreased) by finite heat conduction for $k_\perp/k_z \approx \sqrt{\nu_\perp/\nu_\parallel} < \sqrt{0.1}$ $(>\sqrt{0.1})$ where $\nu_\perp = 0.3b\sqrt{m/2M}\nu_e$ is the perpendicular ion diffusion rate.

[15]Palmadesso, P.J., T.P. Coffey, S.L. Ossakow and K. Papadopoulos, Geophys. Res. Lett., 1, 105, 1974.

[16]Ion Landau damping adds a term $\frac{\omega+\omega_D}{k_z a_i} \Gamma_j(b) \exp\left(-\frac{(\omega-n\Omega)^2}{k_z^2 a_i^2}\right)$ to the right hand side of (6b)-(6d).

The author gratefully acknowledges helpful discussions with Drs. G. Benford, C. Kennel, N. Krall and M. Yamada. This work was supported under NOAA Grant No. 04-5-002-16.

Submitted to Astrophys. J. (1975)

D9

TOWARDS DIFFRACTION-LIMITED SEEING OF DIM ASTRONOMICAL OBJECTS FROM THE EARTH'S SURFACE USING LIDAR

32

Brian G. Cartwright and Ashok J. Gadgil

Department of Physics, University of California,
Berkeley, California 94720

Received_________________

ABSTRACT

It is proposed that LIDAR probing of the atmosphere coupled with active optics be used to eliminate the image-degrading effects of turbulence lying within the bottom few kilometers of the atmosphere above large astronomical telescopes. Available evidence, although scanty, suggests that resolution near the diffraction limit may be possible under some atmospheric conditions at many observatories, and that considerable improvement in resolution would almost always result. If, as is believed likely, image-degrading turbulence at altitudes greater than a few kilometers is confined to a thin layer at the tropopause, a more complex two-pulse, two-correction-plane system can be used to eliminate the residual effects of this layer, thereby providing diffraction-limited seeing over a field-of-view as great as several arc minutes or more for observing any object, no matter how dim.

I. INTRODUCTION

Since the time of Galileo, the resolution of astronomical telescopes has been limited by turbulence-induced spatial and temporal fluctuations of the atmospheric index of refraction, rather than by diffraction at the telescope aperture. The instantaneous image of a point source at infinity obtained by a telescope with an aperture larger than about 10 cm is a speckle pattern arising from multiple-beam interference of the portions of the incoming wave which have traversed different optical path lengths through the atmosphere. On nights of good "seeing" (low image degradation), identical speckle patterns are obtained when one looks at the same point source with the center of one's aperture situated anywhere within a region of about 10 cm diameter, or when one looks at point sources within about 10 μrad of one another. The pattern relaxes within about 10 msec. Systems which can actively correct for atmospheric distortion by analyzing the image obtained from an astronomical object are now being developed by several groups (Muller and Buffington 1974; Feinleib et al. 1974 and Hardy et al. 1974; Bridges 1974). Diffraction-limited performance through the atmosphere over 10 μrad fields-of-view may be obtainable soon for sufficiently bright objects.

Here we propose a technique which may make possible extension of these new methods to dim astronomical objects: the use of LIDAR (Laser RADAR) echoes as an on-line probe of atmospheric conditions.

II. ACTIVE-PUPIL SYSTEMS

Turbulence-compensating systems can be understood by regarding the image aberration as resulting from the effect of an aberration function (phase-shift as a function of position) expressed at one of the optical elements of the telescope, all other optical elements and the air-column being regarded as free of aberration. If this equivalent aberration is detected and if the optical element is "corrected" by suitable mechanical deformation, the aberration may be removed.

Two ways to detect aberration have been proposed to date: shearing interferometric detection (Feinleib et al. 1974) and sharpness-function detection (Muller and Buffington 1974). An interferometric detection scheme which has occurred to us, and which appears very well suited for the system proposed in this paper, in that a single image contains all the information required for correction, is achieved by imaging the deformable optical element through a phase-contrast telescope, operating on the same principles as a phase-contrast microscope. (We have learned privately that R. H. Dicke has also suggested a phase-contrast scheme).

Any scheme using the incoming photons from the astronomical object under observation to measure the aberration function will be limited by photon noise. The aberration function for the atmosphere is known to be band-limited in spatial frequency, so the Whittaker-Shannon sampling theorem is applicable. For an aberration function with a maximum wave number of $2\pi/10\ cm^{-1}$, one needs at least 0.2 samples cm^{-2} of aperture area. An object of visual magnitude $m_v = 12$ viewed in a 3000 Å bandpass through an ideal telescope of 100% optical efficiency and 20% photoelectric efficiency produces in 10 msec less than one photoelectron per 5 cm^2 aperture. Clearly such a weak signal cannot be used to correct for image degradation if the character of that degradation is changing every 10 msec. Thus, $m_v = 12$ is the upper limit on the sensitivity of unaided, active-pupil devices.

Range-gated LIDAR can be used to create a light source at high

altitudes. This light source can then be used to illuminate a phase-contrast telescope. The aberration function observed by the phase-contrast telescope then corrects the astronomical image.
Note that, as in phase-contrast microscopy, where wide-angle reflected skylight is sometimes used as a light source (Bennett 1951), this light source need not be coherent, but must subtend an angle smaller than an isoplanatic patch (about 10 μrad), an angle which is much larger than the Airy resolving power of a large astronomical telescope (about 0.1 μrad). Since we were unable to find a scattering process more effective for the present purposes than Rayleigh scattering, in the following we consider only LIDAR from Rayleigh backscatter.

To calculate the intensity of back-scattered light we have used Rayleigh-scattering and atmospheric attenuation data (Valley 1965) at 3800 Å and 7000 Å to estimate the number of photoelectrons resulting from backscattering into a 10-cm diameter area (one isoplanatic patch) per joule of laser energy for a 1 km range-gated slice of atmosphere. Overall efficiency has been assumed to be 0.005 photoelectrons per photon. The results are shown in Fig. 1. For comparison we note that an existing LIDAR system (Hake et al. 1972) produces about 5 counts at 5890 Å per joule of laser power into a 10-cm diameter receiver from a slice of atmosphere 1-km thick at 20 km. Backscatter intensity as a function of angle from the initial beam direction is shown in Fig. 2, under the assumption of an exponential atmosphere with a scale height of 8 km and a diffraction-limited 10-cm aperture focused at 20 km as a transmitter. Since 10 cm is the size of an isoplanatic region, we might expect the focal point of any aperture larger than 10 cm to be blurred at least as badly as a diffraction-limited 10 cm aperture. The range gate must not be open very long if the echo light source is not to be blurred much beyond the angular extent of an isoplanatic patch. A 1 km slice seems about maximum. For laser pulse repetition rates which are high enough for many laser pulses to occur within the time the atmosphere changes significantly it would be possible to correct the upgoing laser pulse as well. In this case the focused spot could be diffraction-limited. However, at the present time sufficiently powerful lasers with the approximately kHz repetition rates required are not readily available.

III. ECHO ALTITUDE REQUIREMENTS

Since the astronomical object of interest is effectively infinitely distant, the light path from the echo signal will more closely approximate the light path from the astronomical object if the echo is received from as high an altitude as possible. Fig. 3 shows what happens if the turbulence extends over a depth t. If the aperture of the receiver has radius R and the backscattering center is at height H, then the requirement that the backscatter should pass through the same optical path as the light from the astronomical object leads to the requirement that the turbulence layer be thinner than

$$t \leqslant H \frac{10 \text{ cm}}{R} \tag{1}$$

This is a serious limitation, since within the power limitations of contemporary lasers one cannot hope to achieve correction for turbulence at the tropopause, except by rather costly means, e.g., by observing for a fraction of an hour following a rocket-borne sodium vapor release at high altitudes in the desired part of the sky. On the other hand, often seeing may be primarily due to low-lying turbulence.

I. S. Bowen (quoted by Meinel 1960) conducted a qualitative study of the seeing at the 60-inch telescope at Mt. Wilson using Schlieren knife edges to examine double star images. He concluded: "...most of the disturbance to the seeing came from layers within 200 meters of the telescope." Young (1974) summarized the available evidence: "Townsend (1965) ...concluded that two layers with similar temperature fluctuations could explain the observations; his upper layer was near 8 km and ~100-m thick, while the lower layer was the bottom 1 km of atmosphere . . . the upper . . . contributes very little to the seeing. Young (1969) fitted scintillation data with two exponential distributions, one with a scale height of about 8 km and the other with about 1/2 - 1 km scale height but 10-20 times larger temperature fluctuations. Again, the lower layer contributes some 90 percent of the seeing . . ."

Finally, the best guess at present is that image degradation due to turbulence at altitudes above the boundary layer is generally a result of wind shear occurring in a very thin sheet at the tropopause (about 10 km). If this is in fact so, a twin-pulse, two-level correction scheme may be applied: first, a laser pulse is emitted and its echo is studied to determine the aberration resulting from the complex boundary layer of turbulence extending for about 1 kilometer above the telescope aperture. The appropriate correction is then applied to the optics to remove this low-lying seeing. Next, in a time much shorter than 10 msec, a second pulse is emitted, perhaps through optics which also corrects the transmitter for boundary layer turbulence, to probe the structure of the single thin layer at the tropopause. Under the assumption that this layer is thin the argument summarized by equation (1) must be modified, since a single thin layer of turbulence can be removed by a <u>second</u> correction system located at a position relative to the focal plane appropriate to the altitude of the disturbing layer. Thus, one constructs an optical correction scheme approximately conjugate to the atmosphere (to be truly conjugate a large number of correction planes in one-to-one relation to each disturbed layer of the atmosphere would be required). Note that this system would be isoplanatic over the isoplanatic angle associated with the boundary-layer seeing, which is undoubtedly much larger than the observed isoplanatic angle, due to the proximity of the boundary layer to the telescope. Such a system would make possible diffraction-limited seeing over at least several minutes of arc.

IV. CONCLUSION

The altitude structure of the turbulence which produces seeing over observatory sites must be studied before the extent of the contribution of the bottom layer of the atmosphere to seeing can be determined. There is considerable evidence to suggest that by eliminating the effects of this layer an order-of-magnitude improvement in resolution may be possible. If turbulence arising from altitudes above the boundary layer is confined to a thin layer at the tropopause, the residual image degradation resulting from this thin layer can also be removed. Therefore, we believe range-gated LIDAR promises to make possible diffraction-limited seeing from large, ground-based, astronomical telescopes.

REFERENCES

Bennett, A. H. 1951, Phase Microscopy, Principles and Applications (New York: Wiley).

Bridges, W. 1974 (private communication).

Feinleib, J. 1974, Laser Focus, 10, 44.

Hake, R. D., Arnold, D. E., Jackson, D. W., Evans, W. E., Ficklin, B. P., and Long, R. A. 1972, J. Geophys. Res., 77, 6839.

Hardy, J. W., Feinleib, J., and Wayant, J. C. 1974, presented at the "Topical Meeting on Optical Propagation through Turbulence" of the Optical Society of America, 1974 July 9-11, Boulder, Colorado.

Meinel, A. B. 1960, Telescopes, ed. G. P. Kuiper (Chicago: University of Chicago Press), p. 159.

Muller, R. A. and Buffington, A. 1974, J. Opt. Soc. Am., 64, 1200.

Valley, S. L. 1965, Handbook of Geophysics and Space Environments (New York: McGraw-Hill), p. 7-20.

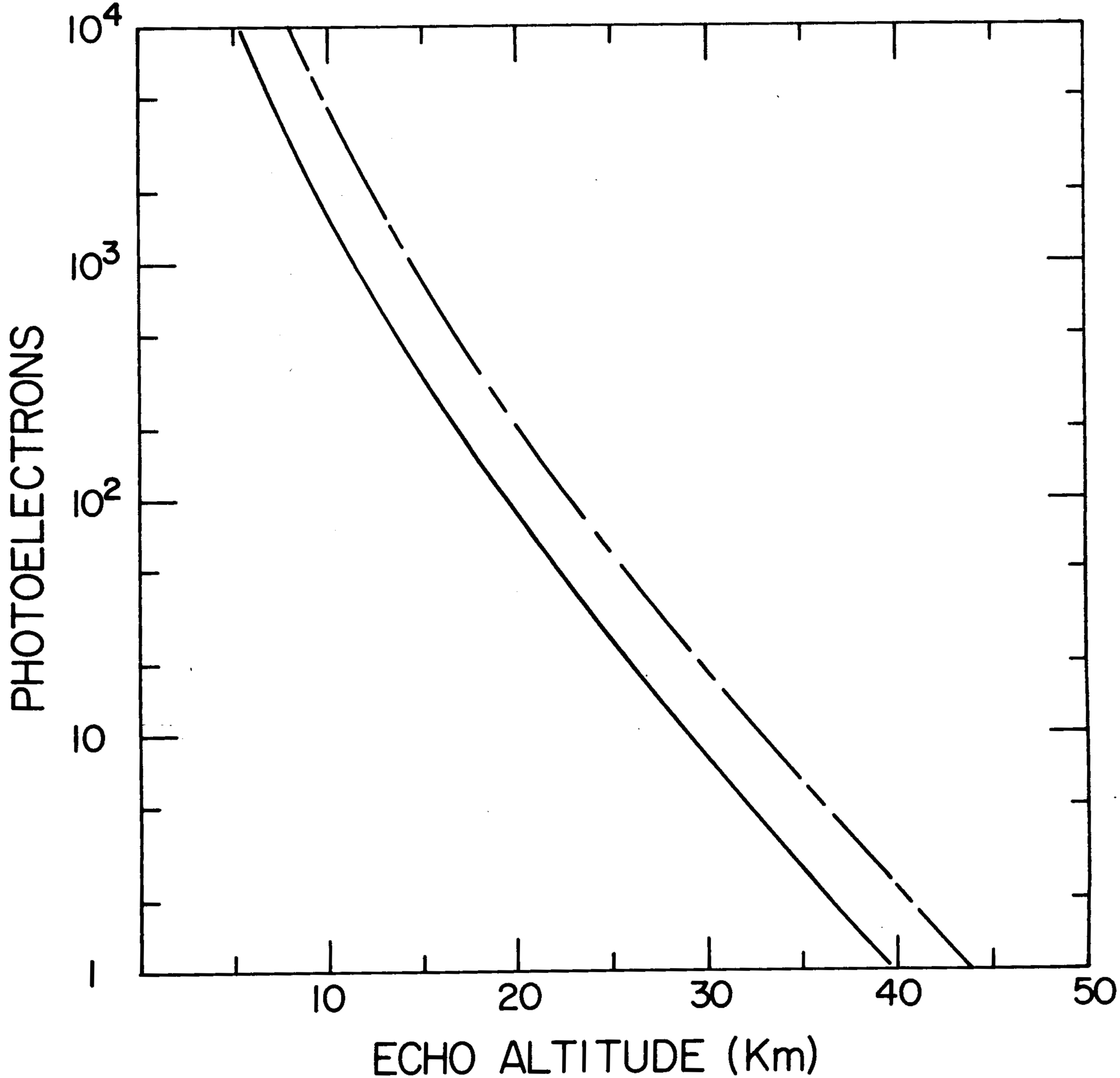
Young, A.T. 1974, Ap. J., 189, 587.

FIGURE CAPTIONS

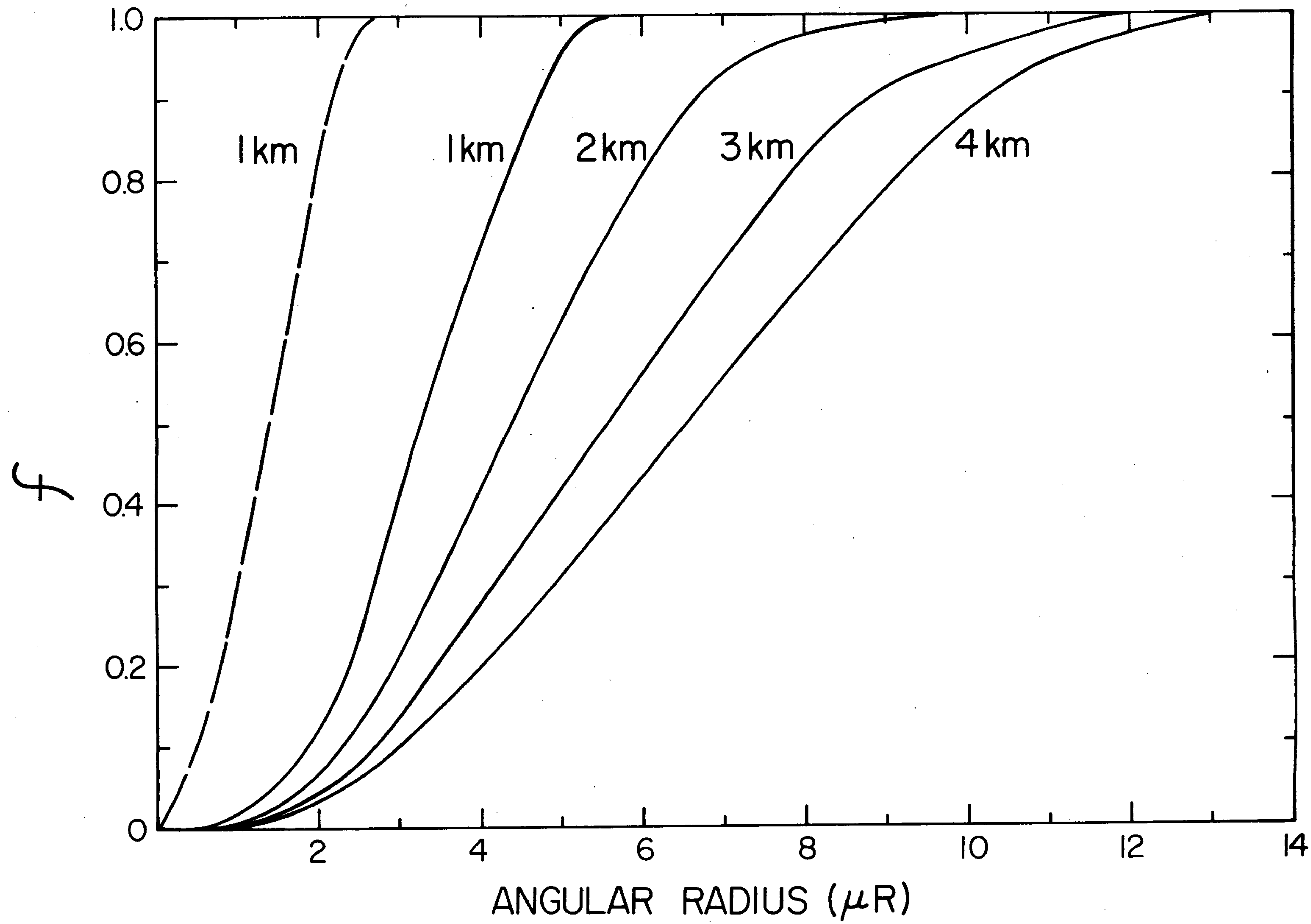
Figure 1. Calculated numbers of counts recorded by a 10-cm diameter LIDAR receiver of efficiency 0.005 photoelectrons per photon resulting from 10-Joule pulses at 3800 Å (broken line) and at 7000 Å (solid line), as a function of echo altitude. The receiver is assumed to be range-gated so as to receive backscattered light from only a 1-km slice of atmosphere at the indicated altitude.

Figure 2. Calculated fraction of the backscattered light contained within a given angular radius as a function of that radius for a range-gated LIDAR focused at 20-km altitude and gated open to receive signals from slices of atmosphere of various thicknesses. A diffraction-limited LIDAR transmitter of 10-cm aperture (solid lines) or 500-cm aperture (broken line) is assumed.

Figure 3. Schematic comparison of the paths followed by light from a LIDAR focused at a height H and by light from a star at infinity. A region of turbulence of thickness t overlies a receiver of aperture 2R.



PHOTOELECTRONS
1
10
10^2
10^3
10^4
10
20
30
40
50
ECHO ALTITUDE (Km)


1.0
0.8
0.6
0.4
0.2
0
f
1km
1km
2km
3km
4km
2
4
6
8
10
12
14
ANGULAR RADIUS (μR)
Figure 2

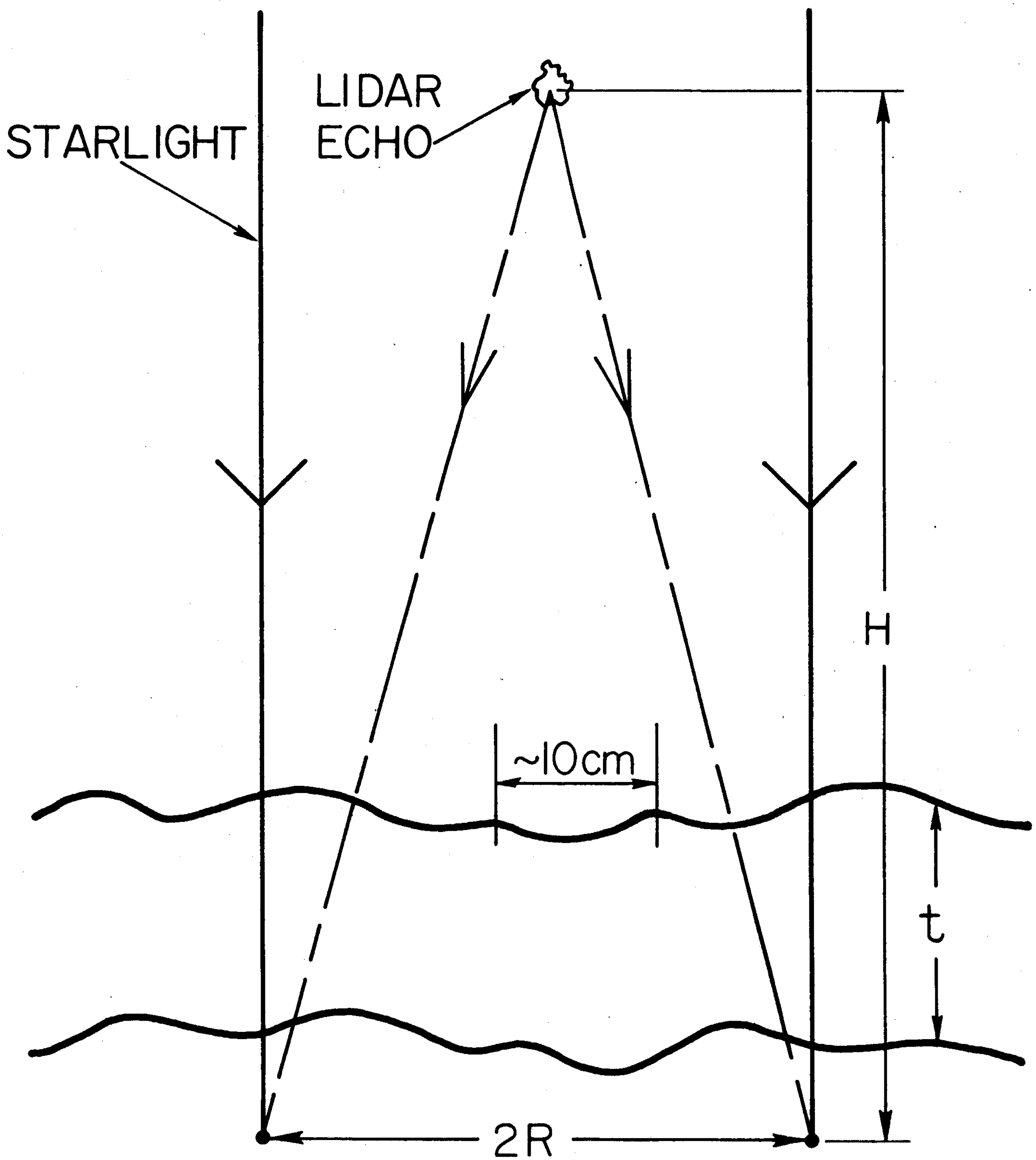


Figure 3